Metal Ions and Protein Folding

Conformational and Functional Interplay

Hugo Miguel Raposo Correia Botelho

Dissertation presented to obtain a PhD degree in Biochemistry
at Instituto de Tecnologia Química e Biológica
Universidade Nova de Lisboa

Supervisor
Cláudio Emanuel Moreira Gomes

Opponents
Teresa J. T. Pinheiro & Peter Faller

Instituto de Tecnologia Química e Biológica
Universidade Nova de Lisboa

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Oeiras, November 2010

Second edition, December 2010

Protein Biochemistry, Folding and Stability Laboratory
Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa
Av. da República (EAN), 2785-572 Oeiras, PORTUGAL
http://www.itqb.unl.pt/pbfs

Cover: Structures of apo (2PRU) and Ca^{2+} and Zn^{2+}-bound (3D10) S100B.

Foreword

This dissertation describes the work performed under the supervision of Cláudio M. Gomes in the Protein Biochemistry Folding and Stability Laboratory at the Instituto de Tecnologia Química e Biológica, from September 2006 to October 2010.

The studies presented herein aim at understanding the role of metal ions in modulating the conformation and stability of proteins. Two types of model systems have been used: iron-sulfur (FeS) proteins having permanently bound FeS cofactors, and the S100 proteins, which bind calcium, zinc and copper ions reversibly.

The first chapter presents the state of the art on fundamental aspects of protein folding. Then, two experimental parts include the studies on iron-sulfur and S100 proteins. Part I includes an overview of the biology of iron-sulfur clusters and their interplay with protein folding and stability followed by three chapters presenting results on the characterization of different FeS proteins (rubredoxin, Rieske ferredoxin and seven iron ferredoxin). Part II first reviews the roles of metal ions as modulators of protein conformations, stability and amyloidogenesis and introduces the S100 proteins. Then, the four following chapters address the modulation of the conformation and stability of S100A2 by metal ions and the newly identified amyloidogenic properties of the S100 family are presented, including the roles of metal ions and the proof of concept for the functional and pathological interplay of S100B and Aβ amyloidogenesis. The last chapter highlights the conclusions of the major findings reported and future perspectives.

Acknowledgements

There were many people who have helped and supported me in different ways during my PhD, to whom I am grateful and want to express my gratitude:

First and foremost, I want to thank my supervisor, Cláudio M. Gomes, for allowing me to pursue exciting research during the last four years, for his joy, foresight, advice and for always trusting in me. By learning from his dynamic and inspiring way of doing science I have grasped how an inquisitive mind can make the difference, broaden my view of science and grown as a scientist.

To Günter Fritz, for the collaboration on the S100 proteins, for great discussions and for always being available to undergo new research projects.

To Ludmilla Morozova-Roche, for sharing her knowledge and for so wonderfully welcoming me in Umeå. Also, to her group, specially Kiran Yanamandra for assisting me in experiments and being such a friendly partner.

To Arnulf Kletzin, for providing Rieske ferredoxin and for discussions on this project.

To the current and past members of the Protein Biochemistry, Folding and Stability Group, specially Bárbara Henriques, Raquel Correia, Sónia Leal, João Rodrigues and Patrícia Faísca, for turning lab relationships in friendship and making being there such a joy. Your companionship will be among the memories from my PhD grant time I will cherish the most.

To my colleague friends Ana Paula Batista, Patrícia Refojo, João Vicente, Sandra Santos, Ana Filipa Pinto and Gabriel Martins for the wonderful shared times, support and the unending good mood. Even though we belonged to different groups, no one who saw us together would say that.

To my family, especially my parents, Victor and Eduarda, for all the love, which does not fade with distance.

To Rita, for her love, support, understanding and so much more.

Fundação para a Ciência e a Tecnologia is acknowledged for financial support, by awarding PhD grant SFRH/BD/31126/2006.

This work was funded by grants POCTI/QUI/45758, PTDC/QUI/70101 and POCI/BIO/58465 from Fundação para a Ciência e a Tecnologia (to Cláudio M. Gomes) and collaborative grant A-15 / 08 from Conselho de Reitores das Universidades Portuguesas and Deutscher Akademischer Austausch Dienst (to Cláudio M. Gomes and Günter Fritz).

Thesis publications

Botelho HM, Koch M, Fritz G & Gomes CM (2009) Metal ions modulate the folding and stability of the tumor suppressor protein S100A2. *FEBS J* 276, 1776-1786

Botelho HM, Leal SS, Veith A, Prosinecki V, Bauer C, Frohlich R, Kletzin A & Gomes CM (2010) Role of a novel disulfide bridge within the all-beta fold of soluble Rieske proteins. *J Biol Inorg Chem* 15, 271-281

Fritz G, **Botelho HM**, Morozova-Roche LA, Gomes CM (2010) Natural and amyloid self-assembly of S100 proteins: structural basis of functional diversity. *FEBS J,* 277(22), 4578-4590

Dissertation abstract

Metal ions are cofactors in about 30% of all proteins, where they fulfill catalytical and structural roles. Due to their unique chemistry and coordination properties they effectively expand the intrinsic polypeptide properties (by participating in catalysis or electron transfer reactions), stabilize protein conformations (like in zinc fingers) and mediate signal transduction (by promoting functionally relevant protein conformational changes). However, metal ions can also exert have deleterious effects in living systems by incorporating in non-native binding sites, promoting aberrant protein aggregation or mediating redox cycling with generation of reactive oxygen and nitrogen species. For this reason, the characterization of the roles of metal ions as modulators of protein conformation and stability provides fundamental knowledge on protein folding properties and is instrumental in establishing the molecular basis of disease. In this thesis we have analyzed protein folding processes using model protein systems incorporating covalently bound metal cofactors – iron-sulfur (FeS) proteins – or where metal ion binding is reversible and associated conformational readjustments – the S100 proteins.

The Rieske [2Fe-2S] domain occurs either as part of electron transfer system in respiratory (bc_1) or photosynthetic (b_6f) complexes or as a component in soluble dioxygenase systems. Despite the functional plasticity, the all-β Rieske fold is highly conserved, being modified by extensions and mutations modulating the cluster redox or pH properties or introducing features like disulfide bridges. In this respect, we have identified a Rieske ferredoxin featuring a cysteine pair present in only a subset of *Thermoprotei* archaea and unrelated to the conserved cysteines forming a disulfide in the vicinity of the respiratory and photosynthetic clusters. By using a combination of thiol colorimetric quantification and the FeS cluster visible and infrared absorption fingerprints we have determined that in the

Acidianus ambivalens protein the cysteines form a disulfide bridge in the as-isolated protein and established the conditions where the disulfide could be selectively reduced (5 mM TCEP) without affecting the overall protein conformation or the redox state of the FeS cluster. Disulfide reduction was found to regulate protein stability (ΔT_m = 9°C), FeS cluster reduction potential (ΔE^0 = +29 mV) and affect the cluster's pH-dependent properties, being a putative regulatory element in this protein.

The folding properties of the hyperthermostable zinc-containing seven iron ([3Fe-4S][4Fe-4S]) ferredoxin from *A. ambivalens* (T_m = 122°C in water at pH 7) have been thoroughly studied. Recently, this has revealed the formation of a molten globule conformation upon thermal unfolding at pH 2.5. This conformation was proposed as a candidate for the template upon which the FeS clusters are assembled during protein biosynthesis. Molten globule formation has been shown to include, zinc and FeS dissociation and tertiary structure changes but a detailed description of the conformational changes occurring is missing. By monitoring the thermal denaturation of ferredoxin at pD 2.5 and 12 using singular value decomposition (SVD) of second derivative FT-IR spectra we determined the thermal denaturation profiles of each secondary structure type of ferredoxin. Further, we identified a spectral component describing ferredoxin's unfolding at acidic and basic conditions. This component was associated with the formation of the molten globule and not of the amorphous unfolded state at basic conditions. The structural state represented by this component undergoes cold unfolding at experimentally accessible temperatures. This is the first description of such event for this protein and constitutes the opportunity of studying cold unfolding of an FeS protein in the absence of external perturbations like co-solvents.

We have built on the knowledge of the folding and stability of rubredoxins by studying the thermal denaturation of the *Desulfovibrio gigas* protein.

Despite its mesophilic origin, this protein exhibited a high stability, with only around ~30% of the FeS clusters disintegrating at 95°C, in an EDTA-insensitive process. By monitoring thermal unfolding using FT-IR coupled with spectral band fitting we have described the partial unfolding occurring at 95°C in terms of the secondary structure changes taking place. We have identified two thermal unfolding regimes. In the 25-60°C range α-helices and β-sheets unfold. Starting at 60°C, turns accumulate and β-sheet content increases, forming around 30 and 50% of the protein's secondary structure at 95°C, respectively. We hypothesize that the highly structured conformation forming at high temperature is correlated with rubredoxin high thermostability.

The S100 proteins are major components in the vertebrate Ca^{2+} signal buffer/transducer network regulating cell cycle, cell growth, differentiation and mobility. S100 proteins are small (10-12 kDa) homo- or heterodimers which bind Ca^{2+} in EF-hand domains. Zn^{2+} and/or Cu^{2+} can also bind elsewhere in some proteins. Ca^{2+} binding induces a conformational change which exposes a protein docking site.

Human S100A2 is a unique family member because it binds Ca^{2+} and Zn^{2+}, accumulates in the nucleus, interacts with p53 in a metal-dependent manner and has been assigned a tumor suppression role. The multiplicity of metal binding sites (2 Ca^{2+} and 2 Zn^{2+} sites per monomer), metal-dependent activation, and S100A2's role in human pathology makes this protein a model to study metal-dependent conformational changes and the eventual tuning of conformational stability. We have examined the role of the Zn^{2+} sites in modulating protein conformation and stability by using S100A2 variants with Zn^{2+} binding site mutations. Circular dichroism analysis has shown that the protein conformation is prone to subtle readjustments upon Ca^{2+} and Zn^{2+} binding, keeping the same α-helical-rich fold. The conformational stability of the S100A2 variants was assessed by CD and FT-IR monitored

urea and thermal unfolding in different metallation states (apo, Ca^{2+}, Zn^{2+}). Thermal denaturation experiments indicated that Zn^{2+} destabilizes and Ca^{2+} stabilizes the protein conformation. These results suggest an opposite role for Ca^{2+} and Zn^{2+} according to which Ca^{2+} activates and stabilizes the protein, and Zn^{2+} inhibits and destabilizes S100A2, a mechanism with possible implications in cancer progression.

Recently, S100 proteins have been shown for the first time to form amyloid fibrils in proteinaceous inclusions associated with pre-carcinogenic inflammatory *foci* in the prostate named *corpora amylacea*. These S100A8/A9 amyloid fibrils could be reconstituted *in vitro* in the presence of Ca^{2+} or Zn^{2+} but not in metal-free conditions. Since S100 family members share sequence and structural homology, we hypothesized that other S100 proteins could also be amyloidogenic. By using the Zyggregator and WALTZ algorithms, equivalent amyloid prone regions were detected in most human S100 proteins. Following this observation, we incubated several S100 family members (S100A2, S100A3, S100A4, S100A6, S100A12 and S100B) in amyloid formation prone conditions (pH 2.5, 57°C) while monitoring amyloid formation by extrinsic thioflavin T (ThT) fluorescence. With the exception of S100A12, all other proteins formed amyloid species. The AFM morphological characterization of the S100 amyloid species revealed that S100A2, S100A6 and S100B formed amyloid oligomers and S100A3 formed amyloid fibrils. For S100B, FT-IR monitored amyloid formation assays revealed that native α-helices and coiled regions convert to β-sheets and turns.

S100 proteins exhibit rich metal binding properties. In addition, S100A6, S100A12, S100B accumulate in *corpora amylacea* in the brain and are overexpressed in neurodegenerative diseases like Alzheimer's, Parkinson's and Amyotrophic Lateral Sclerosis. By using a combination of ThT, FT-IR and AFM monitored amyloidogenesis assays we showed that the S100 amyloid formation kinetics and structure is sensitive to the presence of Ca^{2+}, Zn^{2+} and

Cu^{2+}, major players in the chemical biology of the glutamatergic synapse in the central nervous system. Cu^{2+} promoted S100A12 amyloid formation and had the opposite effect towards S100B. For S100A6, Ca^{2+} completely inhibited amyloidogenesis, a process which could be reverted by adding EDTA. Additionally, Ca^{2+} reverted or alleviated the cytotoxicity of apo S100 amyloids.

Cytotoxicity of S100 amyloids and the modulator role of metal ions is relevant in the framework of the reported involvement of S100 proteins in neurodegenerative processes. Due to the co-accumulation of S100B and Aβ in Alzheimer's disease, we analyzed the possibility of cross-talk between the amyloidogenic processes of both proteins. By carrying out cross-seeding experiments we found that preformed Aβ fibrils seed amyloid formation by S100B. Conversely, S100B oligomers, but not fibrils from the unrelated protein lysozyme, seed Aβ amyloid formation. Overall, our data suggest a cross-talk between Aβ and S100 deposition, cytotoxicity and neurodegeneration, a possibility unexplored in the literature and with relevant biomedical implications.

Resumo da dissertação

Os iões metálicos são cofactores em cerca de 30% de todas as proteínas, desempenhando funções catalíticas e estruturais. Devido às suas propriedades químicas e de coordenação ímpares, alargam as propriedades intrínsecas dos polipéptidos a que se ligam (participando em catálise e reacções de transferência electrónica), estabilizam certas conformações proteicas (como nos dedos de zinco) e servem de mediadores em processos de transdução de sinal (através da promoção de alterações conformacionais funcionalmente relevantes). No entanto, os iões metálicos podem também ter efeitos prejudiciais nos sistemas vivos, ligando-se a locais não nativos, promovendo agregação proteica ou mediando ciclos redox que geram espécies reactivas de oxigénio e azoto. Assim, a caracterização das funções de iões metálicos enquanto modeladores da conformação e estabilidade proteica permite obter conhecimento sobre as propriedades fundamentais do enrolamento e é instrumental na determinação da base molecular de certas doenças. Nesta tese analizámos processos de enrolamento proteico utilizando sistemas modelo que incorporam cofactores metálicos covalentemente ligados – proteínas de ferro-enxofre (FeS) – ou em que a ligação de iões metálicos é reversível e está associada a alterações conformacionais – proteínas S100.

O domínio Rieske [2Fe-2S] ocorre quer como integrante de sistemas de transferência electrónica em complexos respiratórios (bc_1) ou fotossínteticos (b_6f) quer como componente de sistemas dioxigenase solúveis. Apesar da sua plasticidade funcional, a estrutura Rieske β é altamente conservada, sendo modificada por extensões e mutações que modelam as propriedades redox ou dependentes do pH do centro FeS ou que introduzem características como ligações perssulfureto. Neste sentido, identificámos uma ferredoxina Rieske que inclui um par de cisteínas presente apenas num subgrupo de arquea *Thermoprotei* não relacionado com as cisteínas conservadas que formam

uma ligação persulfureto próxima dos centros FeS respiratórios e fotossintéticos. Utilizando uma combinação de quantificação colorimétrica de tióis e os espectros de absorção no visível e infravermelho característicos, determinámos que as cisteínas na proteína de *Acidianus ambivalens* formam um perssulfureto e estabelecemos as condiçõs em que o perssulfureto pode ser reduzido selectivamente (5 mM TCEP) sem afectar a conformação proteica ou o estado redox do centro FeS. A redução do perssulfureto regula a estabilidade proteica (ΔT_m = 9°C), potencial redox do centro FeS (ΔE^0 = +29 mV) e afecta as propriedades dependentes do pH, sendo um possível elemento regulatório nesta proteína.

As propriedades de enrolamento da ferredoxina hipertermostável de sete ferros ([3Fe-4S][4Fe-4S]) com um centro de zinco de *A. ambivalens* (T_m = 122°C em água a pH 7) foram estudadas em detalhe no passado. Mais recentemente, identificou-se a formação de um *molten globule* após desnaturação térmica a pH 2.5. Esta conformação foi sugerida como candidata à estrutura em que os centros FeS são incorporados durante a biossíntese da proteína. Mostrou-se igualmente que a sua formação compreende a dissociação do zinco e do centro FeS bem como alterações da estrutura secundária mas a descrição das alterações conformacionais envolvidas está em falta. Através da monitorização da desnaturação térmica da ferredoxina a pD 2.5 e 12 utilizando a decomposição em valores singulares (SVD) da segunda derivada dos espectros de FT-IR determinámos os perfis de desnaturação térmica para cada tipo de estrutura secundária. Além disso, identificámos um componente espectral que descreve a desnaturação da ferredoxina em condições ácidas e alcalinas. Este componente foi associado com a formação do *molten globule* e não do estado desnaturado amorfo que se forma em condições alcalinas. O estado estrutural representado por este componente sofre desnaturação por frio a temperaturas acessíveis experimentalmente. Esta é a primeira descrição de

tal evento para esta proteína e constitui a oportunidade de estudar a desnaturação por frio de uma proteína FeS na ausência de perturbações externas como co-solventes.

Contribuímos para o conhecimento do enrolamento e estabilidade de rubredoxinas estudando a desnaturação térmica da proteína de *Desulfovibrio gigas*. Apesar da sua origem mesofílica, a proteína é altamente estável, com apenas 30% dos centros FeS a desintegrarem-se a 95°C, um processo insensível ao EDTA. Monitorizando a desnaturação térmica através de FT-IR juntamente com o ajuste de bandas espectrais, descrevemos a desnaturação parcial que ocorre a 95°C em termos de alterações na estrutura secundária. Identificámos dois regimes de desnaturação. Na gama 25-60°C ocorre a desnaturação de hélices α e folhas β. A partir de 60°C, acumulam-se voltas e folhas β, que constituem aproximadamente 30 e 50% da estrutura secundária da proteína a 95°C, respectivamente. Assim, colocamos a hipótese de que a conformação altamente estruturada que se forma a alta temperatura está correlacionada com a estabilidade térmica da rubredoxina.

As proteínas S100 são componentes principais na rede de tamponização e transdução de sinais de Ca^{2+} em vertebrados, regulando o ciclo, crescimento, diferenciação e mobilidade celulares. As proteínas S100 são pequenos (10-12 kDa) homo ou heterodímeros que ligam Ca^{2+} em domínios EF-hand. Zn^{2+} e/ou Cu^{2+} também se podem ligar em locais distintos em algumas proteínas. A ligação de Ca^{2+} induz uma alteração conformacional que expõe uma zona de ancoragem de proteínas.

A proteína S100A2 humana é um membro ímpar desta família porque liga Ca^{2+} e Zn^{2+}, acumula-se no núcleo, interage com o p53 numa forma dependente de metais e tem uma função supressora de tumores. A multiplicidade de locais de ligação a metais (2 locais de Ca^{2+} e Zn^{2+} por monómero), activação dependente de metais e o papel da S100A2 em patologias humanas tornam esta proteína um modelo para o estudo de

alterações conformacionais dependente de metais e o consequente ajustamento da estabilidade. Examinámos o papel dos locais de Zn^{2+} na modelação da conformação e estabilidade proteica utilizando variantes da S100A2 com mutações ao nível dos locais de Zn^{2+}. Uma análise por dicroismo circular mostrou que a conformação desta proteína é propensa a ligeiros reajustamentos em resposta à ligação de Ca^{2+} e Zn^{2+}, mantendo a mesma estrutura rica em hélices α. A estabilidade conformacional das variantes de S100A2 em diferentes estados de metalação (apo, Ca^{2+}, Zn^{2+}) foi quantificada através de desnaturação por temperatura e ureia monitorizada através de CD e FT-IR. As experiências de desnaturação térmica indicaram que o Zn^{2+} desestabiliza e o Ca^{2+} estabiliza a conformação da proteína. Estes resultados sugerem um papel antagónico para o Ca^{2+} e o Zn^{2+} segundo o qual o Ca^{2+} activa e estabiliza a proteína e o Zn^{2+} inibe e desestabiliza a S100A2, um mecanismo com possíveis implicações na progressão do câncro.

Recentemente, foi mostrado pela primeira vez que as proteínas S100 formam fibras amilóides em inclusões proteicas associadas com focos pré-inflamatórios na próstata denominadas *corpora amylacea*. Estas fibras amilóides de S100A8/A9 podem ser reconstituidas *in vitro* na presença de Ca^{2+} ou Zn^{2+} mas não em condições livres de metais. Dado que as proteínas S100 partilham homologia estrutural e de sequência, colocámos a hipótese de que outras proteínas S100 também pudessem ser amiloidogénicas. Utilizando os algoritmos Zyggregator e WALTZ, foram identificadas regiões com propensão amiloidogénica na maioria das proteínas S100 humanas. Na sequência desta observação, incubámos vários membros da família S100 (S100A2, S100A3, S100A4, S100A6, S100A12 e S100B) em condições que favorecem a formação de amilóide (pH 2.5, 57°C) enquanto monitorizávamos a formação de amilóide através da fluorescência extrínseca de tioflavina T (ThT). Com a excepção da S100A12, todas as outras proteínas formaram espécies amilóides. A caracterização morfológica das espécies amilóides

S100 por AFM revelou que a S100A2, S100A6 e S100B formaram oligómeros amilóides e que a S100A3 formou fibras amilóides. A formação de amilóides de S100B monitorizada por FT-IR revelou que hélices α e regiões desestruturadas nativas se convertem em folhas β e voltas.

As proteínas S100 possuem ricas propriedades de ligação a metais. Além disso, a S100A6, S100A12 e S100B acumulam-se em *corpora amylacea* no cérebro e são sobre-expressas em doenças neurodegenerativas como Alzheimer, Parkinson e Esclerose Lateral Amiotrófica (ALS). Através da utilização da combinação de ensaios de amiloidogénese monitorizados por fluorescência de ThT, FT-IR e AFM, mostrámos que a cinética de formação de amilóides S100 e a respectiva estrutura é sensível à presença de Ca^{2+}, Zn^{2+} e Cu^{2+}, elementos principais da química biológica da sinapse glutamatérgica no sistema nervoso central. O Cu^{2+} promoveu a formação de amilóide S100A12 e teve o efeito oposto relativamente à S100B. Relativamente à S100A6, o Ca^{2+} inibiu completamente a amiloidogénese, um processo que pôde ser revertido pela adição de EDTA. Para além disto, o Ca^{2+} reverteu ou diminuiu a citotoxicidade dos amilóides S100 apo.

A citotoxicidade dos amilóides S100 e o papel modelador dos iões metálicos é relevante no âmbito do anteriormente reportado envolvimento das proteínas S100 em processos neurodegenerativos. Devido à co-acumulação de S100B e Aβ na doença de Alzheimer, analizámos a possibilidade de inter-relação entre os processos amiloidogénicos de ambas. Através de experiências de *cross-seeding* descobrimos que fibras Aβ pré-formadas funcionam como sementes para a formação de amilódes S100B. Por outro lado, oligómeros S100B, mas não fibras da proteína não relacionada lisozima, são sementes na formação de amilóide Aβ. Globalmente, os resultados sugerem uma inter-relação entre a deposição de Aβ e proteínas S100, citotoxicidade e neurodegeneração, uma possibilidade

anteriormente não explorada na literatura e com relevantes implicações biomédicas.

Abbreviations

AD	Alzheimer's disease
AFM	Atomic force microscopy
ALS	Amyotrophic lateral sclerosis
ANS	1-anilino-8-naphthalene-sulfonic acid
APP	Amyloid precursor protein
ATR	Attenuated total reflectance
Aβ	Amyloid β
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
CD	Circular dichroism
C_m	Midpoint denaturant concentration
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E^0	Standard reduction potential
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EPR	Electron paramagnetic resonance
ESI-MS	Electrospray ionization mass spectrometry
Fd	Ferredoxin
FeS	Iron-sulfur
FRET	Förster resonance energy transfer
FT-IR	Fourier transform infrared spectroscopy
GuHCl	Guanidinium hydrochloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
K_d	Dissociation constant
KPi	Potassium phosphate
MCT	Mercury cadmium telluride
MES	2-(N-morpholino)ethanesulfonic acid
mRNA	Messenger ribonucleic acid
NAD^+	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
PD	Parkinson's disease
PrP	Prion protein
RAGE	Receptor for advanced glycation endproducts
Rd	Rubredoxin
SDS	Sodium dodecyl sulfate
SOD	Superoxide dismutase
SVD	Singular value decomposition
TCEP	Tris(2-carboxyethyl)phosphine
ThT	Thioflavin T
T_m	Midpoint denaturation temperature
T-ramp	Temperature ramp
TSE	Transition state ensemble
USE	Unfolded state ensemble
UV	Ultraviolet
WST	Water soluble tetrazolium
ΔG	Gibbs free energy variation
ΔΔG	ΔG variation
ε	Extinction coefficient

Contents

General introduction

1. Protein folding

1.1. The protein folding problem

The biological activity of proteins frequently depends on the ability of the polypeptide to acquire a very defined and unique tree-dimensional structure. This is not surprising as catalysis, signal transduction, ligand binding and molecular interactions – the main functions of proteins – all require a stringent spatial arrangement of the polypeptide chain: the native structure. This structure is attained through protein folding, the physical process by which a polypeptide folds into its characteristic and functional three-dimensional structure from a random coil [1]. Protein folding is a fundamental process in biology due to the high dependence of practically all biological processes on the protein machinery. However, even in the cellular environment, a fraction of all synthesized proteins fail to fold into the native structure [2]. This inability may bring about severe biological consequences such as the so-called misfolding diseases [3] which are associated with degradation prone-conformations (such as in cystic fibrosis [4-5]), misfolded protein with an aberrant activity (like in phenylketonuria [6] or fatty acid metabolism disorders [7]) or protein deposition in the form of insoluble amyloid fibrils (characteristic of neurodegenerative diseases as Alzheimer's [8], Parkinson's [9] or Huntington's [10]).

Since Anfinsen's seminal work on ribonuclease [1] we know that proteins can fold spontaneously to their lowest free energy structure without the intervention of additional entities. This means that all information specifying the three dimensional structure of a protein is encoded solely on its amino acid sequence. This observation has profound implications in biology. It means that the sole knowledge of the amino acid sequence in a protein should be enough for unequivocally determining its folded structure in any given medium and, consequently, its interaction and catalytic properties and the folding mechanism. However, this task is not trivial due to the many degrees of freedom arising from the polymeric nature of proteins, the

heterogeneity of amino acid side chains, the multiplicity of interactions stabilizing the folded conformation and an incomplete mechanistic understanding of the protein folding process. Accordingly, this task is referred as the Protein Folding Problem, from which three main questions arise [11]:

1. What is the energetics stabilizing folded proteins?
2. What is the folding mechanism?
3. Can the protein structure be predicted from the sole knowledge of the amino acid sequence?

Topics 1 and 2 are discussed in sections 1.2 and 1.3, respectively. The large structural biology data currently available highlights the complexity of the protein fold problem [12]. Although the interactions determining the native state are the same for all proteins and this determines a restricted number of secondary structure elements, there is an enormous number of possible spatial arrangements (Figure 1.1). Moreover, protein size and oligomerization status offer further diversification potential. However, proteins occupy a discrete portion of the conformational space: the number of protein folds – the topological arrangements of secondary structure elements – is restricted (Figure 1.1).

Despite the inherent complexity of protein folding, in recent years, computational modeling coupled with single molecule studies have contributed with significant progress to a mechanistic description of protein folding process at the molecular level. In this framework, the experimental contribution has been limited because most experimental determinations rely in ensemble measurements which intrinsically average fine structural information. Useful strategies are single molecule FRET and force spectroscopy. Computational studies have yielded more quantitative information. The great challenge in modeling protein folding is two-fold: first, the large amount of atoms in the simulation with the concomitant large

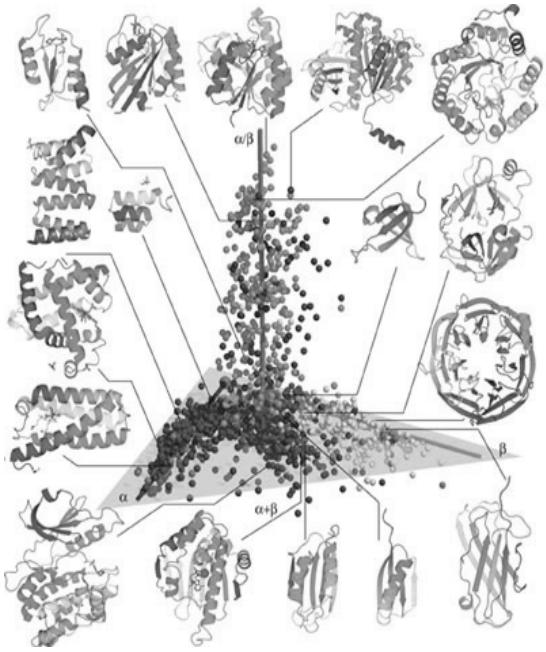


Figure 1.1 – Three-dimensional representation of the protein conformational space. Each sphere represents a protein fold family among compact globular proteins. The structures were placed in space according to their pair-wise structure alignments. The structures cluster according to SCOP protein fold classes. Representative structures are shown to highlight the structure variability and different secondary structure arrangements using the same restricted number of basic elements. From [12].

conformational space; second, the timescale of protein folding requires highly computer intensive calculations. Earlier studies based in minimalistic Gō-type models have been complemented by more realistic off-lattice models. Increasing computer power, including distributed computing projects such as Rosetta@home of the Foldit game [13], allow routine full atomistic molecular dynamics and *ab initio* structure prediction studies. Currently folding simulations have reached the millisecond timescale [14]. The quality of the protein structure prediction algorithms is assessed periodically in the CASP competition – critical assessment of protein structure prediction – against experimentally determined protein structures. Progress is significant and automatic protein structure prediction is a possibility in a not too distant future.

1.2. Stabilization of the folded state

1.2.1. Overall view

The folding status (folded/unfolded) of proteins is dictated by the balance between two energetic parameters. On the one hand, the unfolded state has high conformational entropy because proteins are linear polymers of tens to hundreds of amino acid residues; on the other side, the multiplicity of interactions involving side chains, the backbone and the solvent constitute an enthalpic gradient towards the folded state. Folding is then driven by the enthalpy of interactions forming while interacting groups are brought into proximity. The thermodynamic forces responsible for the folding of proteins to their native conformations were first described by Kauzmann [15]: hydrophobic interactions, hydrogen bonds and electrostatic interactions [16-17]. These interactions are sometimes referred as "weak interactions" in protein folding. This is because of their small nominal magnitude. However, the very significant number of individual interactions in folded proteins accounts for a high overall energetic term which keeps the native protein structure. In some cases, the folded conformation is further stabilized by additional contributions arising from disulfide bridges, oligomerization, cofactor binding or post-translational modifications. Together, all these molecular interactions contribute to building up all levels of protein structure.

1.2.2. The hydrophobic effect

The major contribution to folding comes from the hydrophobic effect, arising from the backbone, aromatic and aliphatic side chains [15, 18]. In the cell, proteins fold in a highly dielectric aqueous environment [19]. Water forms entropically costly "water cages" around hydrophobic areas [20], driving hydrophobic self-association which minimizes the entropic penalty. The outcome is that native protein structures frequently exhibit a highly

packed hydrophobic core surrounded by a polar shell facing the solvent [21]. The hydrophobic contribution to protein folding can be determined through mutagenesis [22] or by the free energy change of transferring an amino acid side chain from water to a hydrophobic solute [23]. Average stabilization values are 1.3 kcal/mol per buried methyl group or 1-2 kcal/mol per aromatic ring interaction [22]. The vast hydrophobic areas buried during protein folding account for the stabilizing importance of the hydrophobic effect.

1.2.3. Hydrogen bonds

Hydrogen bonds are also important stabilizing features in protein structures. The backbone and several amino acid side chains have hydrogen bond donor and acceptor groups. Further, the solvent can also be included in hydrogen bonding patterns. The contribution of hydrogen bonds to the overall protein stabilization is very relevant. About two thirds of all residues form peptide hydrogen bonds in the native state [24]. It is the involvement of the backbone in repetitive interactions that drives secondary structure formation [25]. The stabilization by hydrogen bond averages to 1.3 kcal/mol [26-27]. For geometric reasons, the same hydrogen bond donor or acceptor can participate in multiple hydrogen bonds, creating hydrogen bond networks.

1.2.4. Ion pairs

The most individually stabilizing interaction in protein folding is the electrostatic interaction between charged amino acid side chains, called ion pairs. Ion pairs are usually located at the protein surface, being favored by the aqueous polar environment. A single ion pair can contribute with up to 3-5 kcal/mol to stability [28]. This value is close to the typical stability of the native versus the unfolded state (ΔG_{N-U}), meaning that in some proteins breaking a single ion pair may result in protein unfolding. Despite their

individual magnitude, the average number of ion pairs in proteins is small. Consequently, they are not major driving forces in protein stabilization [18].

1.2.5. Covalent modifications

Disulfide bridges are cross-links which effectively restrict the conformational space, decreasing protein conformational entropy and increasing protein stability [29-30]. The most significant effects occur with disulfides bridging distant polypeptide regions. For avoiding the formation of aberrant disulfides – which form off-pathway folding species – organisms have developed specialized enzymatic mechanisms [31]. Prokaryotes produce a family of disulfide bond proteins (Dsb) which assist oxidative protein folding – with the concomitant cysteine thiol oxidation – and correct aberrant disulfides [32]. In eukaryotes, the endoplasmic reticulum is the oxidative cellular compartment where disulfide formation is regulated by protein disulfide isomerase (PDI) and Ero1 [33].

Similarly to disulfide bridges, the binding of cofactors also cross-links different parts of the polypeptide and favors folding. Proteins can bind a broad range of organic or inorganic cofactors in a covalent or non covalent manner. For example, zinc finger domains can only fold to their native structure in the presence of the metal ion [34]. Cytochrome *c* requires its covalently bound heme for folding and function [35-36]. Calcium is frequently associated with conformational fine-tuning [37]. Also, the contribution from post-translational modifications such as glycosylation, phosphorylation and lysine methylation is relevant well [16].

1.2.6. Two state protein folding thermodynamics

The native state is then stabilized by the combination of multiple mutually supportive weak interactions. The outcome is that the folded conformations of proteins are only marginally stable even under the most favorable conditions [38-39]: typically ~5-15 kcal/mol, in the range of just a few non

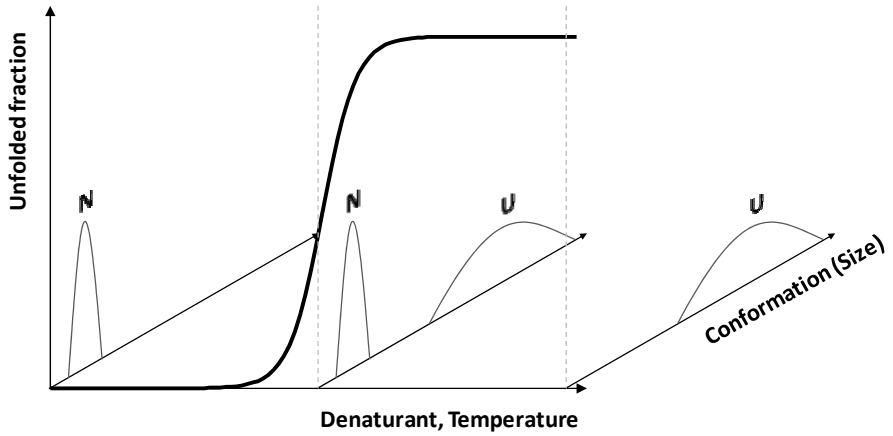


Figure 1.2 – Two state unfolding. Denaturing conditions (chemical denaturant or temperature) unfold a two state folder in a cooperative manner. At any condition, only the native (N) and/or unfolded (U) state are detected and interconvert. The native state is a unique state while the unfolded state ensemble groups a broader set of conformations.

covalent interactions. No energetic component is preponderant in the overall stabilization. Even the simplest protein folding energetic potentials must include multiple energetic components to reproduce protein folding [17].

Small single domain proteins usually unfold reversibly [1] and according to a simple two state process. This means that the only conformations detectable are the native and the unfolded state. These two interconvert and no intermediately folded state is populated at detectable levels (Figure 1.2). Due to the cooperative nature of protein folding, such a protein is perturbed by an increasingly denaturing environment, the conformation initially changes very little [40]. Then, for a limited range of conditions, the protein unfolds completely. In this simple unfolding scenario, the same unfolding profile will be obtained independently of the technique used to probe it [40].

The thermodynamics of two state reversible unfolding is characterized as a function of the native (N)/unfolded (U) equilibrium constant, K_{eq}:

$$K_{eq} = \frac{[N]}{[U]}$$

from which the free energy variation is obtained

$$\Delta G_{N-U} = G_N - G_U = -RT\ ln(K_{N-U})$$

where R and T are the gas constant and absolute temperature, respectively.

The thermodynamic characterization of lysozyme unfolding at pH 7 carried out by Privalov and co-workers [41] illustrates the complex relationship between the thermodynamic parameters that determine protein stability. The opposing effects of denatured state conformational entropy (TS^0) and native stabilizing interactions (H^0) individually comprise several hundred kcal/mol. However, both contributions have similar temperature trends and compensate each other, implying marginal protein stability. The Gibbs free energy of folding exhibits a temperature dependence ($\Delta G_{N-U}(T)$)

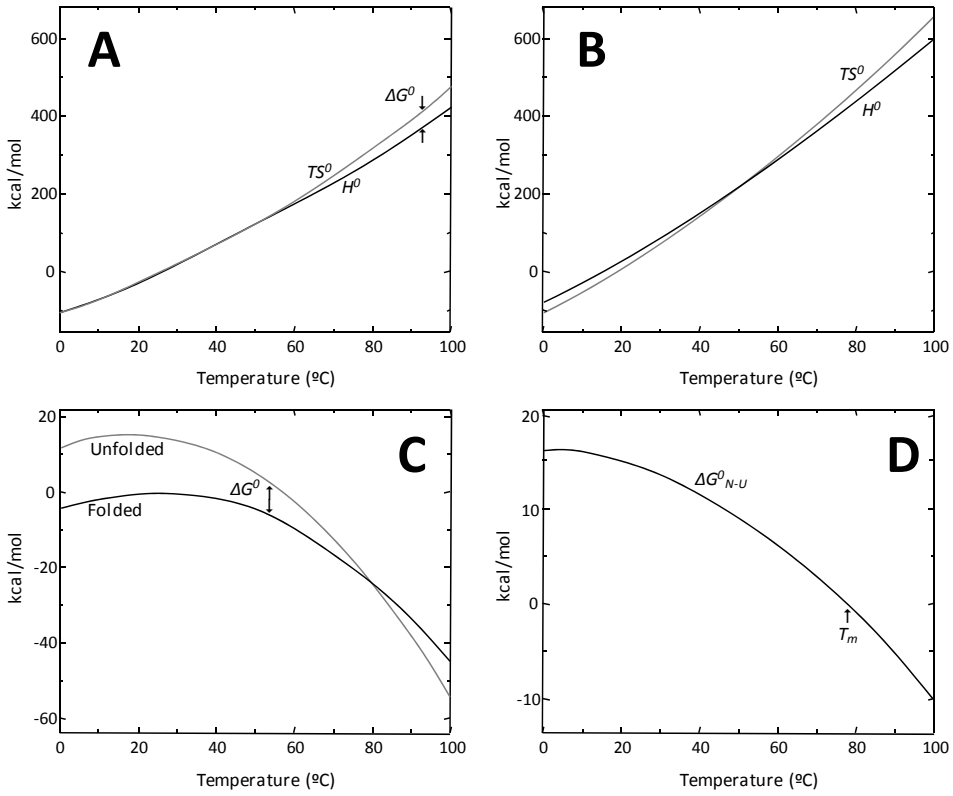


Figure 1.3 – Temperature-dependent thermodynamic folding parameters for lysozyme at pH 7. The enthalpic (*H*) and entropic (*TS*) contributions are very large and of similar magnitude for the folded (A) and unfolded (B) states. The most populated state is the lowest energy one (C). The energetic difference between the folded and unfolded states is small and results in marginal protein stability (D). At the point where ΔG_{N-U} is null, there is equal amount of protein in both the folded and unfolded states. This temperature is referred to as the midpoint denaturation or melting temperature (T_m). From [41].

expressed by the Gibbs-Helmholtz equation (Figure 1.3). If the heat capacity (ΔCp_{N-U}) and apparent enthalpy (ΔH_{N-U}^{app}) difference between the native and the unfolded states is assumed temperature-independent, the equation has the following formula:

$$\Delta G_{N-U}(T) = \Delta H_{N-U}^{app}\left(1 - \frac{T}{T_m}\right) + \Delta Cp_{N-U}(T - T_m) - T\ln\left(\frac{T}{T_m}\right)$$

The Gibbs-Helmholtz equation expresses the fact that ΔG_{N-U} is maximal at some temperature and that there are two temperatures where it equals zero (*i. e.* proteins can be denatured by heat or by cold). Heat denaturation occurs because of the enthalpic compensation of stabilizing interactions by temperature. Cold denaturation is due to the decreased hydrophobic contribution at lower temperatures [42]. Experimentally, usually only heat denaturation is observed because the cold denaturation temperature is frequently below the water melting temperature. Nevertheless, this has been observed for a few proteins [43].

1.2.7. Engineering protein stability

Several studies have shown that the stability of naturally occurring proteins is not optimized and can be improved by mutagenesis [39, 44], sometimes paralleled with function enhancement [44-47]. However, in the extreme case, increased stability may generate a highly compact and stiff, non functional native state [39]. This suggests that marginal protein stability is the evolutionary response to the need to achieve the native state (*i. e.* to have a minimum energy state [1]) while preserving (*i.*) the protein dynamics required for catalysis, conformational plasticity, protein recognition and proteolytic susceptibility [39], a key element in cellular protein turnover [48]; and (*ii.*) an energetic buffer which allows protein evolution [49] (Figure 1.4).

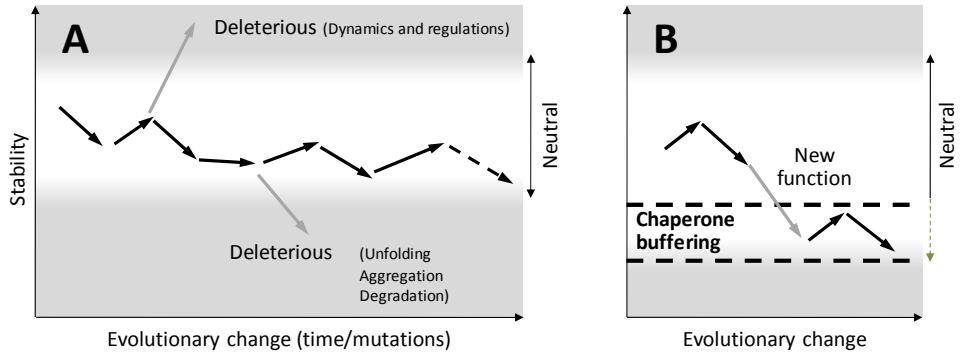


Figure 1.4 – Protein evolution. (A) Along evolutionary time, proteins accumulate mutations which impact folding and stability. The marginal conformational stability of proteins implies that only a subset of all allowed mutations are tolerated. Mutations outside this neutral zone may are deleterious because they affect the protein's ability to fold (destabilizing mutations) or to be functionally dynamic (over-stabilizing mutations). (B) Molecular chaperones rescue the folding of misfolded species and create an enhanced "neutral area" from where new functions may arise. From [49].

1.3. Protein folding models

Since the discovery that the protein native structure is determined by its amino acid sequence, much effort has been made in elucidating the mechanistics of protein folding. The first milestone in this field was set by Cyrus Levinthal in the late 1960s [50]. He elegantly pointed out that if a small 99 peptide bond protein would randomly sample just three rotamers of each of its 198 phi and psi angles in the typical timescale of molecular rotations – 1 picosecond - it would take 10^{75} years to fold – much more than the age of the universe. However, proteins fold in a biologically relevant timescale (typically in the microsecond to second timescales [51]). This became known as Levinthal's Paradox. It encloses a fundamental aspect of protein folding: the range of accessible conformations during folding is restricted. For solving this paradox, Levinthal postulated the existence of a pathway consisting of a well-defined and restricted sequence of protein conformational changes bridging the unfolded to the folded state [52] – the sequential model of protein folding. However, no mechanistic description to describe what the

folding pathway could be was put forward. Several models about protein folding have however been conceived since then (Figure 1.5).

In the late 1970s Levitt had already identified the caveat in Levinthal's paradox: the search for the native state is not *unbiased*. He noted that protein folding could be reproduced *in silico* as a random search if native-like interactions are considered to be on average more stabilizing than non-native ones. Thus, according to this random search model [53] the folding pathway is analogous to solving a jigsaw puzzle, where the order of the steps is not relevant but the end result is always the same native state [54]. More recently, the kinetics and thermodynamics of the folding of the villin headpiece domain has been successfully modeled using this formalism [55].

Kim and Baldwin proposed the framework model [53]. This model considers that secondary and tertiary structural elements form independently of one another. Secondary structure elements build up in the beginning of the folding reaction and progressively assemble into the native tertiary structure.

The hydrophobic collapse model [56] hypothesizes that folding initiates by a rapid polypeptide collapse driven by hydrophobic side chain self-association, resulting in the formation of an intermediate state devoid of secondary structure and with non-native tertiary structure, *i. e.* a molten-globule-like state. The structural reorganization of the molten globule from the restricted conformational space available – accounting for the folding energy barrier – would give rise to secondary structure and the native fold. The validation of this model is complex because it relies on the structural characterization of early, short-lived folding intermediates. Nevertheless, it appropriately describes the folding kinetics of globular proteins such as myoglobin [57], α-lactalbumin [58], barstar [59] and staphylococcal nuclease [60].

Karplus and Weaver proposed the diffusion-collision model [61-62]. In this case, the folding protein is considered to be composed of several independent marginally stable secondary structure microdomains each one exhibiting fast conformational dynamics. During folding individual microdomains collide and eventually adhere and coalesce to give rise to the native tertiary structure.

The classical nucleation model proposes that the folding reaction is guided by the formation of a marginally stable nucleus containing correct secondary and tertiary structural elements. The place of nucleation in the folding reaction was matter of dispute: Wetlaufer proposed that nucleation would comprise the folding onset [63] and Baldwin proposed that nucleation would be the limiting factor [64]. In any case, the nucleus templates formation of further structure around it, restricts the available conformational space and speeds folding without implying the existence of folding intermediate states [65]. In the 1990s Monte Carlo simulations of a lattice model by Shakhnovich and co-workers [66] supported that nucleation limits folding. Once the nucleus is assembled, the native conformation is promptly formed. This discovery permanently associated the study of protein folding with that of the transition state. The nucleation model describes the folding of two state folders but fails to describe folding processes where intermediates accumulate [67].

In the 1990s, Fersht and co-workers established that the folding of chymotrypsin inhibitor 2 (CI2) – a 64 residue two state folder – could only be described by a new folding model: the nucleation-condensation model [68-70]. This model postulates the existence of a marginally stable nucleus composed mainly by adjacent residues early in folding. The rate limiting step is the eventual stabilization of this small nucleus by long range interactions. This extended nucleus is not formed in the transition state but represents the best formed part of the structure at this point [71]. Nucleation is coupled to

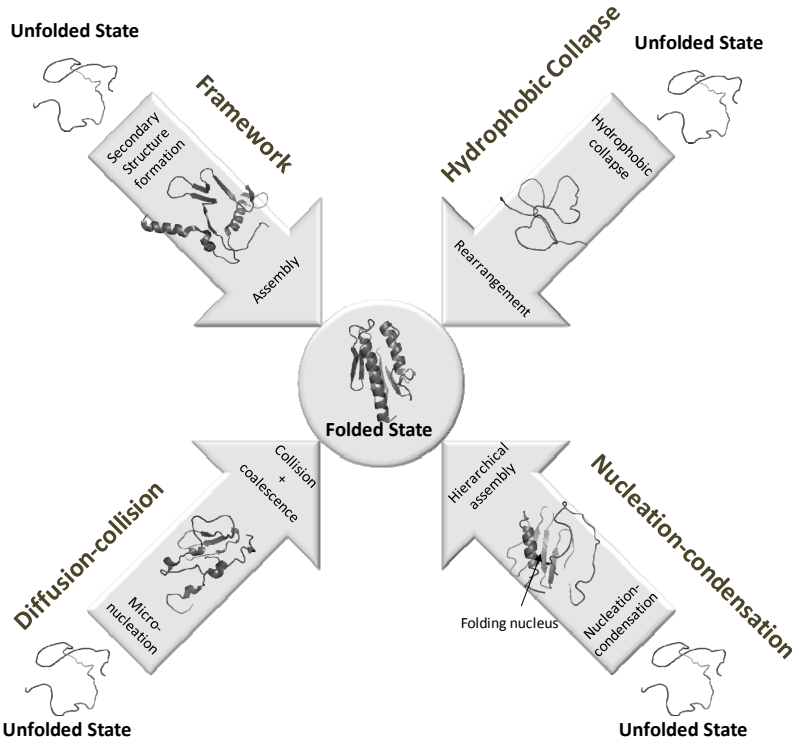


Figure 1.5 - Protein folding models.

condensation of tertiary structure around it. Since the proposal of this model, other proteins were shown to adopt compatible folding kinetics [72].

1.4. Protein folding kinetics

Protein folding is a structural event. Nevertheless, it can be appropriately described by standard chemical reaction kinetics theory. The starting state – the unfolded state ensemble – is regarded as a "reagent" and the end state – the native state – as a "product". The reaction mechanism, energetics and kinetics can be described by monitoring the interconversion of the two entities, possibly through some intermediate states. Like a reaction, protein folding is a diffusive process, *i. e.* an unfolded polypeptide will spontaneously acquire the most stable accessible conformational state given the solvent conditions through a stochastic search.

The classical way of studying protein folding *in vitro* is through temperature-, pressure-, acid- or denaturant-induced renaturation [71]. The folded and unfolded states frequently exhibit distinct spectroscopic properties. Coupling renaturation (by manual or stopped flow mixing, temperature or pressure jumps) to spectroscopic detection allows monitoring the transition between the two states and describing folding [73]. Information about folding can also be obtained by studying the analogous unfolding transition. A summary of the main techniques used in studying protein folding – not just the kinetics – is given in Table 1.1.

The conformational search inherent to protein folding implies that only the lowest energy states (typically the native state) are unique and well defined, all others (*e. g.* the unfolded and transition states) being in fact an

Table 1.1 – Experimental techniques useful for the study of protein folding. From [73].

Technique	Timescale	Information content
Intrinsic tryptophan fluorescence	≥ ns	Tryptophan environment
Far UV CD	≥ μs	Secondary structure content
Near UV CD	≥ μs	Aromatic residue packing
Raman spectroscopy	≥ μs	Solvent accessibility, aromatic residues' conformation
Infrared spectroscopy	≥ ns	Secondary structure content
ANS binding	≥ μs	Hydrophobic surface exposure
FRET	≥ ps	Molecular ruler
Fluorescence correlation spectroscopy	≥ ps	Diffusion, size and shape
Fluorescence anisotropy	≥ μs	Shape and size
Small-angle X-ray scattering	≥ μs	Radius of gyration
Absorbance	≥ ns	Chromophore environment
Real-time NMR	≥ min	Structure
Native-state hydrogen exchange	h	Global stability
Pulsed H/D exchange by NMR	≥ ms	Solvent accessibility
Pulsed H/D exchange by ESI-MS	≥ ms	Solvent accessibility
NMR relaxation	~ms	Denatured state structure, conformational changes
Protein engineering	Probe dependent	Residue-specific information

ensemble of distinct states. Then, the aforementioned states are accurately named the unfolded state ensemble (USE) and the transition state ensemble (TSE). The reason for the occasional language simplification is due to the fact that these states are most frequently experimentally accessible though their statistically averaged properties. The TSE is defined as the set of conformations such that folding trajectories starting from each one of them have a 50% probability of either reaching to the folded state before unfolding and reaching the unfolded state before folding [74].

Protein folding is a cooperative process [75]: the establishment of native or native-like contacts facilitates further interactions. This speeds folding [76] and stabilizes the native state once it is formed, as full unfolding requires the cumulative loss of the interaction in the cooperative network. In the extreme case, as it is frequently found in small globular proteins, cooperativity leads to two state folding [77] where the only detected populated states are the native and the unfolded ones, which interchange between themselves during folding. In this situation, protein folding kinetics is monoexponential. The surprisingly fast protein folding rates are achieved through a combination of factors:

1. Presence of residual structure in the fully unfolded state which restricts the accessible conformational space during folding;
2. The folding pathway comprises metastable intermediates which work like hubs directing folding to the native state [78];
3. The funneled energy landscapes biases folding to the native state.

There would be no native state search problem if each of the amino acids could find its native conformation independently of the others or if only nearest-neighbor interactions were involved. This would reduce the protein folding problem to an analog of the helix-coil transition. What makes the search problem difficult is that long-range interactions are involved. It is the presence of long-range effects that make the folding transition cooperative

(pseudo-first order), an essential element of the stability and kinetics of proteins [75, 79].

In the framework of the success of the nucleation formalism in describing the folding of small globular proteins, the folding transition state has been routinely probed by a protein engineering method named Φ-value analysis [71]. It consists in analyzing the effect of single point mutations in the protein's folding behavior and determining a thermodynamic parameter – the Φ-value – defined as

$$\Phi = \frac{-RT\ ln\left(\frac{k_{mut}}{k_{wt}}\right)}{\Delta\Delta G_{N-D}} \approx \frac{\Delta\Delta G_{\ddagger-D}}{\Delta\Delta G_{N-D}}$$

where k_{mut} and k_{wt} are the folding rates of the mutant and wild type proteins, respectively, $\Delta\Delta G_{N-D}$ is the change in folding free energy upon mutation and $\Delta\Delta G_{\ddagger-D}$ is the change in activation free energy upon mutation (Figure 1.6). The approximation is valid for non-disruptive (*i. e.* having a small perturbation) mutations. The Φ-value then accounts for the degree of native-like interactions at the mutated position in the transition state. A unity value means that the perturbation extent is the same in the native and transition states, indicating that the specified residue has all native interactions formed in the native state. A value close to zero means that the residue has the same interactions in the denatured and in the transition state. Structural interpretations of Φ-values have been routinely put forward, identifying residues with Φ-values equal to 1 or 0 as being fully structured or unstructured in the transition state, respectively. However, controversy exists about the interpretation of Φ-values outside these values. Fractional Φ-values have been interpreted as the indication of alternative folding pathways [71] and non-classical Φ-values (Φ<0, Φ>1) have been interpreted as reflecting the differential perturbation of secondary and tertiary structural elements [80].

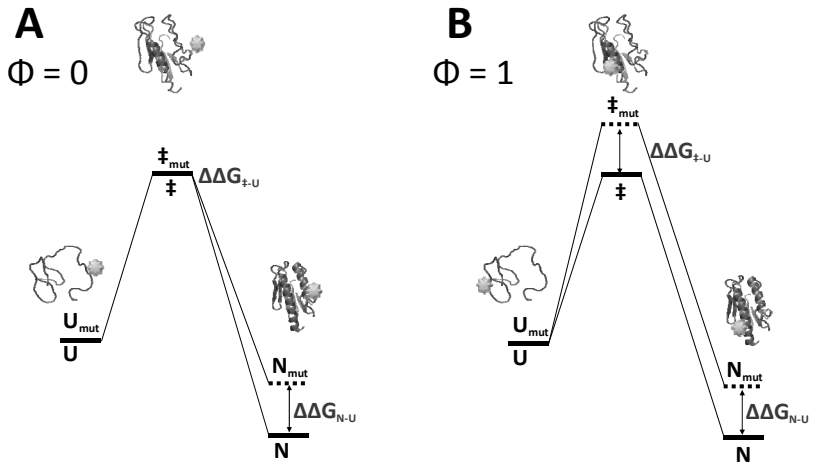


Figure 1.6 – Φ-value analysis of protein folding. The putative folding energy diagrams for two point mutants are sketched. (A) The mutation does not perturb the interactions and, consequently, the energetics of the selected amino acid residue in the unfolded nor in the transition states. Nevertheless, the mutation is destabilizing. The Φ-value is then zero. (B) The mutation affects a residue which is part of the folding nucleus. The transition state is then destabilized in the same amount as the native state and the Φ-value equals 1. From [71].

Proteins which were selected by evolution to fold into a biologically relevant structure are able to satisfy the local structure propensity of amino acid residues (dictated by their respective rotamers) and intramolecular interactions (responsible for stabilizing the native state) without contradicting each other. Of course this can only occur in native or native-like structures, guiding the folding process. This situation is called the "minimal frustration" of protein folding [81-82].

So, protein folding is a highly complex mechanism. Part of the complexity comes from the definition of the starting point, the unfolded state. Unlike the native state, the unfolded state is a broad collection of conformationally distinct states with high entropy – the denatured state ensemble. It contributes with structural heterogeneity to the starting state, and determines a lot of the random search for the native state. The inherent stochasticity of protein folding results in different parts of the protein being folded at different times in the folding reaction, almost independently of each

other. These folding units forming in one single cooperative step are called "foldons" [81-82]. Since every intermediate *en route* to the native state has a different conformation, it also has a specific energy. The stochastic fluctuation between sequential folding intermediates gives rise to a distorted, rugged funnel, populated by a myriad of energetic basins (or kinetic traps, following a kinetic analysis) which stabilize misfolded intermediates and slow down folding. To overcome these traps, local or global unfolding must occur. Some basins can be transposed within k_BT. Deeper ones can become conformational dead-ends from which the protein cannot be rescued without external assistance (*e. g.* by molecular chaperones), despite the favorable thermodynamics. Folding intermediates have been detected experimentally for many systems including hen lysozyme [83], cytochrome *c* [84] among others.

The extreme case of kinetically-controlled folding is achieved for proteins whose native state – determined by its biological activity – is not the lowest energy one but rather a metastable higher energy state trapped by a large activation energy barrier. This is possible for proteins which include prosegments in the newly-synthesized polypeptide which are subsequently proteolytically excised. Such post-translational processing occurs in insulin [85], α-lytic protease [86], pepsin [87] and the serpins [88].

1.5. Protein folding energetics – the landscape view

A rigorous solution to Levinthal's Paradox, describing both the thermodynamics and the kinetics of folding, comes from the statistical analysis of the energetics of the folding mechanism. This consists in considering the energy of the folding protein sampled from the whole set of conformers it can assume – a multidimensional "conformational space". Proteins can then be viewed as folding according to the dependence of conformation on energetics – the "energy landscape". The unfolded state ensemble collapses and reconfigures into multiple inter-convertible

conformations. The exact sequence of conformational transitions is stochastically determined – like the jigsaw puzzle concept [54]. However, since the Gibbs free energy (ΔG) of each conformation is different and the native state has the lowest energy – Anfinsen's thermodynamic hypothesis – the energy landscape is biased towards the native state. The formation of stabilizing interactions precludes the sampling of competing conformations greatly restricting the accessible conformational space. On the other hand, incorrect local folds tend to be eliminated by more stable conformations arising in the fluctuations inherent to this diffusive process [82, 89]. The energetic gradient also implies that not all conformations are equally likely in the folding pathway, solving Levinthal's paradox, and that native-like intermediates may act as hubs guiding folding [90]. This situation has been pictured by Onuchic and co-workers as a funneled energy landscape [91].

The folding funnel determines protein folding kinetics and thermodynamic properties. Its shape is a function of the medium composition. Events such as solvent change, external perturbations (*e. g.* temperature, pressure), protein association or ligand binding change the shape of the folding funnel and, ultimately, protein conformation. This is the thermodynamic basis for protein allostery, protein association, protein conformational changes and protein unfolding.

The bottleneck inherent to the folding funnel acts to speed folding. In small globular proteins this produces a single exponential folding kinetics (Figure 1.7) [77]. However, the many degrees of freedom in protein folding produce a folding funnel which is not smooth but rather rugged, with many energetic – *i. e.* kinetic – traps. These local minima illustrate the situation where non-native interactions may be locally stabilizing but are not compatible with achieving the global energy minimum. In other words, the folding energy landscape incorporates some "energetic frustration" (Figure 1.8D) [92-94]. However, the fast and cooperative folding of most natural

proteins shows that energetic traps have been smoothed by evolution rendering naturally occurring proteins "minimally frustrated" [89, 95].

Growing evidence supports the hypothesis that the folding energy landscape may in fact include additional folding funnels accounting for misfolded species [97]. Due to their simplicity, the folding of naturally occurring small single domain proteins is frequently associated with smooth energy landscapes, giving rise to single exponential folding kinetics. In the framework of landscape theory it is possible to envisage an extreme situation where a determined polypeptide sequence folds to the native state without encountering significant energetic barriers throughout the folding funnel (Figure 1.8B), a situation termed "downhill folding" [98]. In such a scenario, folding kinetics is non-exponential and all conformers are *en route* to the native state, making them accessible to experimental study, including structural characterization. Downhill folding is a sort of limit situation for real folding [98] as it can be captured theoretically by coarse-grained Gō-type models [99]. Here, the polypeptide is modeled as a string-and-bead structure where amino acid residues – the beads – can only occupy defined positions in a three dimensional cubic lattice and interact if they occupy contiguous positions in space.

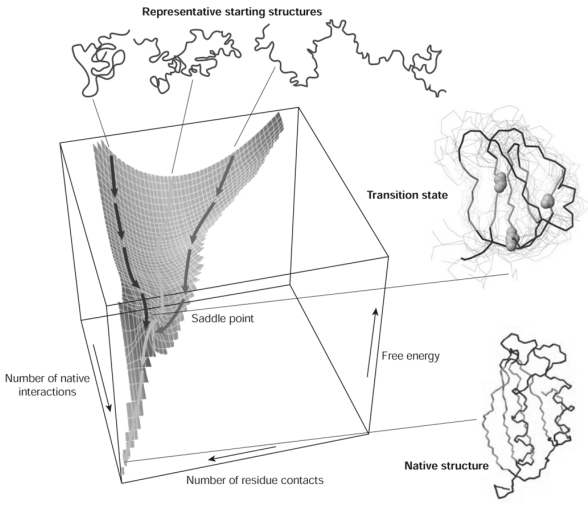


Figure 1.7 - A schematic folding funnel of acylphosphatase. The great conformational space accessible to the unfolded state is progressively restricted as the protein lowers its free energy through folding. The energetic gradient progressively biases the conformational search to the native state, although multiple folding pathways are still accessible. Once the folding transition state barrier, depicted as a saddle point, is overcome by the formation of the folding nucleus (depicted as yellow spheres in the structure) folding is fast and productive. From [96].

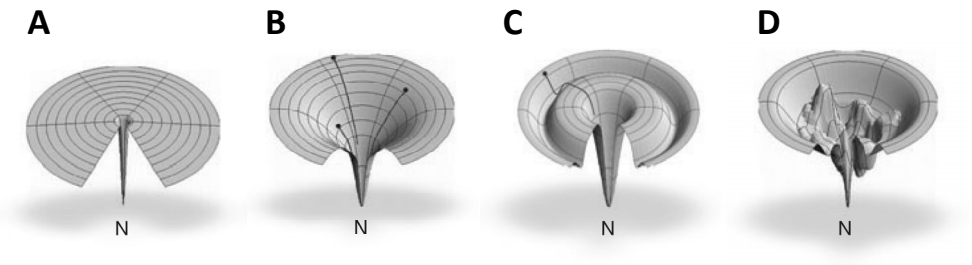


Figure 1.8 - Protein folding funnels. (A) Levinthal's funnel. All folding pathways are equally likely and there is no folding cooperativity in the search for the native state. (B) Downhill folding funnel. Smooth energy landscape without kinetic traps. (C) Folding funnel for a protein with an obligatory on-pathway folding intermediate. (D) Highly frustrated folding funnel, with many kinetic traps and alternative folding routes. From [100].

1.6. Conformational states – a unified view of protein folding

The highly complex interaction network in proteins coupled with the large conformational space accessible results in a series of conformational states being accessible to proteins. Unlike the classical view of protein folding whereby a polypeptide is synthesized in the ribosome in the unfolded state and eventually attains the monomeric or oligomeric native state, several intermediate or off-pathway states do exist and have been shown to be relevant in the biologic context of protein folding. Partially folded intermediates may expose hydrophobic patches which drive self-aggregation. Misfolded or aggregated protein may be recognized by cellular quality control systems and degraded. It is believed that, given appropriate conditions, every protein has the potential of acquiring the β-sheet rich aggregation-prone amyloid conformation, which may polymerize into fibrils (Figure 1.9) [96].

The exception to this panorama are proteins which, notwithstanding the possibility of having local persistent structure, the majority of the polypeptide chain is in a random coil-like conformation: the so-called intrinsically unstructured proteins or natively unfolded proteins. These proteins are not able to acquire a compact native state because of having a high charge density. These proteins are not without functional significance.

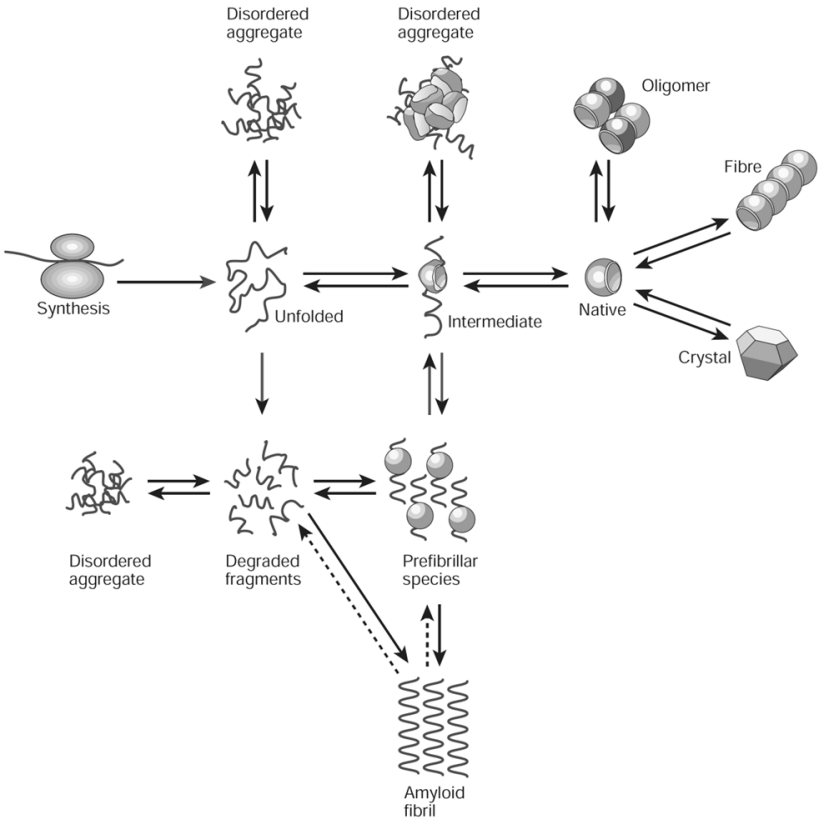


Figure 1.9 – Protein conformational states. From the moment a polypeptide is synthesized in the ribosome, it can adopt a multitude of conformations apart from the unfolded and native states. Intermediately folded states may be aggregation- or degradation-prone. The β-sheet rich amyloid state is an aggregated state which is thought to be accessible to all proteins. Artificially, proteins can be superconcentrated and forced to adopt a crystalline state. From [96].

Their high conformational dynamics renders them promiscuous protein binders. So, these proteins fulfill mainly protein network integration roles, working like hubs binding proteins from different signaling pathways and contributing to signal transduction [101-104].

1.7. Protein folding in the cell

In the biologic context of protein folding – newly synthesized proteins exiting the ribosome – multiple extrinsic factors affect folding. First, protein synthesis is vectorial – the N-terminus is synthesized prior than the C-

terminus – and the synthesis rate is slow (~2-4 amino acid residues per second in eukaryotic systems [105]), likely constituting the rate-limiting factor in protein folding *in vivo*. Since protein folding is dictated by long range interactions which may involve interactions between the N- and C-terminus, especially in multidomain proteins (*e. g.* [106]), and shape the folding energy landscape [107] non-native interactions may occur during protein synthesis. Secondly, protein biosynthesis occurs in the context of densely packed polysomes (Figure 1.10), further favoring interaction of aggregation-prone solvent-exposed hydrophobic residues. To minimize improper interactions, ribosomes are oriented around the mRNA molecule in a pseudohelical arrangement which minimizes the interactions between vicinal nascent chains [108]. The ribosome exit channel is hydrophilic [109] and favors co-translational folding [110] of at least α-helical elements [111-113]. Cell environment is highly crowded, with macromolecular concentrations reaching as much as 400 mg/ml [114-116]. The excluded volume effect favors protein misfolding and aggregation [116-118]. However, it also biases the folding landscape towards compact conformations [119], restricting the accessible conformational space and speeding folding in some situations [116-117, 120-121]. As a response, cells accumulate compatible solutes, also known as osmolytes [122], sometimes up to molar concentrations, which favor protein hydration and, consequently, hydrophobic burial and folding. To cope with the challenges of protein folding, biological systems have evolved a specialized protein quality control machinery (Figure 1.11) aimed at aiding or correcting folding – the molecular chaperones – or, ultimately, degrading terminally misfolded or aggregated proteins [123].

Molecular chaperones recognize and reversibly bind nascent polypeptides or non-native protein conformations (hydrophobic amino acid stretches [125] which are buried in the folded state or unstructured protein regions),

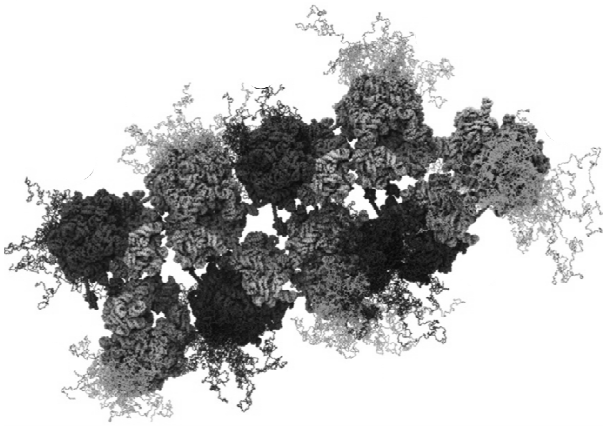


Figure 1.10 – Polysomes. Several ribosomes transcribing a single mRNA molecule (in the center). The newly synthesized polypeptides are depicted as red or green threads on the exterior of the polysome. The high protein density favors aggregation of the misfolded polypeptides. To overcome this problem, polypeptides exit the ribosomes at opposing faces. From [108].

reduce aggregation propensity and assist folding. They do not integrate the final native structure. Some proteins are facultative or obligatory [125] chaperone substrates needing their assistance to be rescued from intermediate states [67] and fold on biologically relevant timescales [126]. Molecular chaperones act by either stabilizing the nascent chain in the ribosome or by downstream action [127-129]. The first group of chaperones stabilizes cytosolic proteins by binding to the large ribosome subunit and eventually to hydrophobic protein sequences, maintaining an extended non-aggregated protein state until enough polypeptide has been synthesized for productive folding to occur. This group includes the bacterial trigger factor, the eukaryotic RAC (ribosome-associated complex, an Hsp70 system) and the archaeal and eukaryotic NAC (nascent chain-associated complex).

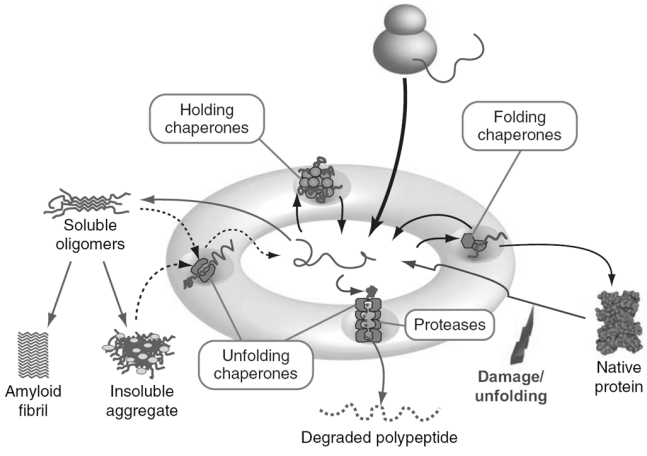


Figure 1.11 - Protein quality control systems. As the newly synthesized protein exits the ribosome, surveillance mechanisms act to detect, correct and ultimately degrade misfolded species. Holding chaperones prevent protein aggregation; folding chaperones assist the folding and unfolding chaperones act on aggregated or misfolded protein to promote further folding cycles or degradation by proteases. From [124].

Downstream-acting chaperones do not bind the ribosome and assist co- or post-translational folding. These include members of the Hsp70 family (DnaK in bacteria, Hsc70 in higher eukaryotes [125]) and further downstream systems like the chaperonins (Hsp60 family: GroEL in bacteria, thermosome in archaea and TRiC/CCT in eukarya [130-132]) and Hsp90 [133]. Broad-specificity chaperones such as members of the Hsp70 or chaperonin families primarily recognize hydrophobic amino acid side chains and promote their folding through ATP-dependent binding-release-rebinding cycles. If a protein is terminally misfolded and cannot be rescued by chaperone action, the polypeptide is degraded. In archaea, eukaryotes and some bacteria, the most important proteolytic system is the ubiquitin-proteasome pathway [134]. It involves tagging the misfolded protein with multiple ubiquitin chains through the sequential action of three ligases and subsequent proteolytic degradation by the proteasome, a multiprotein complex.

1.8. References

1. Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230
2. Yewdell JW (2005) Serendipity strikes twice: the discovery and rediscovery of defective ribosomal products (DRiPS). *Cell Mol Biol (Noisy-le-grand)* **51**, 635-641
3. Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* **75**, 333-366
4. Qu BH, Strickland E & Thomas PJ (1997) Cystic fibrosis: a disease of altered protein folding. *J Bioenerg Biomembr* **29**, 483-490
5. Cheung JC & Deber CM (2008) Misfolding of the cystic fibrosis transmembrane conductance regulator and disease. *Biochemistry* **47**, 1465-1473
6. Gersting SW, Kemter KF, Staudigl M, Messing DD, Danecka MK, Lagler FB, Sommerhoff CP, Roscher AA & Muntau AC (2008) Loss of function in phenylketonuria is caused by impaired molecular motions and conformational instability. *Am J Hum Genet* **83**, 5-17
7. Gregersen N, Bross P & Andresen BS (2004) Genetic defects in fatty acid beta-oxidation and acyl-CoA dehydrogenases. Molecular pathogenesis and genotype-phenotype relationships. *Eur J Biochem* **271**, 470-482
8. Blennow K, de Leon MJ & Zetterberg H (2006) Alzheimer's disease. *Lancet* **368**, 387-403
9. Uversky VN & Eliezer D (2009) Biophysics of Parkinson's disease: structure and aggregation of alpha-synuclein. *Curr Protein Pept Sci* **10**, 483-499
10. Shastry BS (2003) Neurodegenerative disorders of protein aggregation. *Neurochem Int* **43**, 1-7
11. Dill KA, Ozkan SB, Shell MS & Weikl TR (2008) The protein folding problem. *Annu Rev Biophys* **37**, 289-316
12. Hou J, Sims GE, Zhang C & Kim SH (2003) A global representation of the protein fold space. *Proc Natl Acad Sci U S A* **100**, 2386-2390
13. Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popovic Z & players F (2010) Predicting protein structures with a multiplayer online game. *Nature* **466**, 756-760

14. Voelz VA, Bowman GR, Beauchamp K & Pande VS (2010) Molecular simulation of ab initio protein folding for a millisecond folder NTL9(1-39). *J Am Chem Soc* **132**, 1526-1528
15. Kauzmann W (1959) Some factors in the interpretation of protein denaturation. *Adv Protein Chem* **14**, 1-63
16. Vieille C & Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* **65**, 1-43
17. Yang JS, Chen WW, Skolnick J & Shakhnovich EI (2007) All-atom ab initio folding of a diverse set of proteins. *Structure* **15**, 53-63
18. Dill KA (1990) Dominant forces in protein folding. *Biochemistry* **29**, 7133-7155
19. Floriano WB & Nascimento MAC (2004) Dielectric constant and density of water as a function of pressure at constant temperature. *Braz J Phys* **34**, 38-41
20. Teeter MM (1984) Water structure of a hydrophobic protein at atomic resolution: Pentagon rings of water molecules in crystals of crambin. *Proc Natl Acad Sci U S A* **81**, 6014-6018
21. Munson M, Balasubramanian S, Fleming KG, Nagi AD, O'Brien R, Sturtevant JM & Regan L (1996) What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Protein Sci* **5**, 1584-1593
22. Burley SK & Petsko GA (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23-28
23. Dill KA (1990) The meaning of hydrophobicity. *Science* **250**, 297-298
24. Stickle DF, Presta LG, Dill KA & Rose GD (1992) Hydrogen bonding in globular proteins. *J Mol Biol* **226**, 1143-1159
25. Marqusee S, Robbins VH & Baldwin RL (1989) Unusually stable helix formation in short alanine-based peptides. *Proc Natl Acad Sci U S A* **86**, 5286-5290
26. Shirley BA, Stanssens P, Hahn U & Pace CN (1992) Contribution of hydrogen bonding to the conformational stability of ribonuclease T1. *Biochemistry* **31**, 725-732
27. Myers JK & Pace CN (1996) Hydrogen bonding stabilizes globular proteins. *Biophys J* **71**, 2033-2039
28. Anderson DE, Becktel WJ & Dahlquist FW (1990) pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**, 2403-2408
29. Matsumura M, Signor G & Matthews BW (1989) Substantial increase of protein stability by multiple disulphide bonds. *Nature* **342**, 291-293
30. Meyer J, Clay MD, Johnson MK, Stubna A, Munck E, Higgins C & Wittung-Stafshede P (2002) A hyperthermophilic plant-type [2Fe-2S] ferredoxin from Aquifex aeolicus is stabilized by a disulfide bond. *Biochemistry* **41**, 3096-3108
31. Gruber CW, Cemazar M, Heras B, Martin JL & Craik DJ (2006) Protein disulfide isomerase: the structure of oxidative folding. *Trends Biochem Sci* **31**, 455-464
32. Collet JF & Bardwell JC (2002) Oxidative protein folding in bacteria. *Mol Microbiol* **44**, 1-8
33. Tu BP & Weissman JS (2004) Oxidative protein folding in eukaryotes: mechanisms and consequences. *J Cell Biol* **164**, 341-346
34. Frankel AD, Berg JM & Pabo CO (1987) Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc Natl Acad Sci U S A* **84**, 4841-4845
35. Fisher WR, Taniuchi H & Anfinsen CB (1973) On the role of heme in the formation of the structure of cytochrome c. *J Biol Chem* **248**, 3188-3195
36. Stellwagen E, Rysavy R & Babul G (1972) The conformation of horse heart apocytochrome c. *J Biol Chem* **247**, 8074-8077
37. Yamniuk AP, Silver DM, Anderson KL, Martin SR & Vogel HJ (2007) Domain stability and metal-induced folding of calcium- and integrin-binding protein 1. *Biochemistry* **46**, 7088-7098
38. Bloom JD, Raval A & Wilke CO (2007) Thermodynamics of neutral protein evolution. *Genetics* **175**, 255-266
39. DePristo MA, Weinreich DM & Hartl DL (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet* **6**, 678-687
40. Nojima H, Ikai A, Oshima T & Noda H (1977) Reversible thermal unfolding of thermostable phosphoglycerate kinase. Thermostability associated with mean zero enthalpy change. *J Mol Biol* **116**, 429-442

41. Pfeil W & Privalov PL (1976) Thermodynamic investigations of proteins. III. Thermodynamic description of lysozyme. *Biophys Chem* **4**, 41-50
42. Creighton TE (1992) Stability of the folded conformation. In *Proteins: Structures and Molecular Properties*. W. H. Freeman.
43. Martin SR, Esposito V, De Los Rios P, Pastore A & Temussi PA (2008) Cold denaturation of yeast frataxin offers the clue to understand the effect of alcohols on protein stability. *J Am Chem Soc* **130**, 9963-9970
44. Foit L, Morgan GJ, Kern MJ, Steimer LR, von Hacht AA, Titchmarsh J, Warriner SL, Radford SE & Bardwell JC (2009) Optimizing protein stability in vivo. *Mol Cell* **36**, 861-871
45. Shoichet BK, Baase WA, Kuroki R & Matthews BW (1995) A relationship between protein stability and protein function. *Proc Natl Acad Sci U S A* **92**, 452-456
46. Martin A, Kather I & Schmid FX (2002) Origins of the high stability of an in vitro-selected cold-shock protein. *J Mol Biol* **318**, 1341-1349
47. Wang X, Minasov G & Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol* **320**, 85-95
48. Forloni G, Iussich S, Awan T, Colombo L, Angeretti N, Girola L, Bertani I, Poli G, Caramelli M, Grazia Bruzzone M, et al. (2002) Tetracyclines affect prion infectivity. *Proc Natl Acad Sci U S A* **99**, 10849-10854
49. Tokuriki N & Tawfik DS (2009) Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol* **19**, 596-604
50. Levinthal C (1968). In *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois* (Debrunner P, Tsibris JCM & Münck E, eds), pp. 22-24. University of Illinois Press, Urbana, Monticello, IL.
51. Klepeis JL, Lindorff-Larsen K, Dror RO & Shaw DE (2009) Long-timescale molecular dynamics simulations of protein structure and function. *Curr Opin Struct Biol* **19**, 120-127
52. Levinthal C (1968) Are there pathways for protein folding? *J Chim Physique* **65**, 44
53. Kim PS & Baldwin RL (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu Rev Biochem* **51**, 459-489
54. Harrison SC & Durbin R (1985) Is there a single pathway for the folding of a polypeptide chain? *Proc Natl Acad Sci U S A* **82**, 4028-4030
55. Sullivan DC & Kuntz* ID (2002) Protein Folding as Biased Conformational Diffusion. *J Phys Chem B* **106**, 3255-3262
56. Baldwin RL (1989) How does protein folding get started? *Trends Biochem Sci* **14**, 291-294
57. Gilmanshin R, Dyer RB & Callender RH (1997) Structural heterogeneity of the various forms of apomyoglobin: implications for protein folding. *Protein Sci* **6**, 2134-2142
58. Arai M & Kuwajima K (1996) Rapid formation of a molten globule intermediate in refolding of alpha-lactalbumin. *Fold Des* **1**, 275-287
59. Agashe VR, Shastry MC & Udgaonkar JB (1995) Initial hydrophobic collapse in the folding of barstar. *Nature* **377**, 754-757
60. Vidugiris GJ, Markley JL & Royer CA (1995) Evidence for a molten globule-like transition state in protein folding from determination of activation volumes. *Biochemistry* **34**, 4909-4912
61. Karplus M & Weaver DL (1976) Protein-folding dynamics. *Nature* **260**, 404-406
62. Bashford D, Cohen FE, Karplus M, Kuntz ID & Weaver DL (1988) Diffusion-collision model for the folding kinetics of myoglobin. *Proteins* **4**, 211-227
63. Wetlaufer DB (1973) Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc Natl Acad Sci U S A* **70**, 697-701
64. Baldwin RL & Rose GD (1999) Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem Sci* **24**, 26-33
65. Faísca PFN (2009) The nucleation mechanism of protein folding: a survey of computer simulation studies. *J Phys: Condens Matter* **21**
66. Abkevich VI, Gutin AM & Shakhnovich EI (1994) Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry* **33**, 10026-10036
67. Brockwell DJ & Radford SE (2007) Intermediates: ubiquitous species on folding energy landscapes? *Curr Opin Struct Biol* **17**, 30-37

68. Fersht AR (1995) Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc Natl Acad Sci U S A* **92**, 10869-10873
69. Fersht AR (1997) Nucleation mechanisms in protein folding. *Curr Opin Struct Biol* **7**, 3-9
70. Itzhaki LS, Otzen DE & Fersht AR (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J Mol Biol* **254**, 260-288
71. Fersht A (1998) Kinetics of protein folding. In *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*. Freeman, San Francisco.
72. Nolting B & Agard DA (2008) How general is the nucleation-condensation mechanism? *Proteins* **73**, 754-764
73. Bartlett AI & Radford SE (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat Struct Mol Biol* **16**, 582-588
74. Mirny L & Shakhnovich E (2001) Protein folding theory: from lattice to all-atom models. *Annu Rev Biophys Biomol Struct* **30**, 361-396
75. Privalov PL (1989) Thermodynamic problems of protein structure. *Annu Rev Biophys Biophys Chem* **18**, 47-69
76. Faisca PF & Plaxco KW (2006) Cooperativity and the origins of rapid, single-exponential kinetics in protein folding. *Protein Sci* **15**, 1608-1618
77. Jackson SE & Fersht AR (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-10435
78. Karplus M (1997) The Levinthal paradox: yesterday and today. *Fold Des* **2**, S69-75
79. Karplus M & Weaver DL (1994) Protein folding dynamics: the diffusion-collision model and experimental data. *Protein Sci* **3**, 650-668
80. Weikl TR & Dill KA (2007) Transition-states in protein folding kinetics: the structural interpretation of Phi values. *J Mol Biol* **365**, 1578-1586
81. Panchenko AR, Luthey-Schulten Z, Cole R & Wolynes PG (1997) The foldon universe: a survey of structural similarity and self-recognition of independently folding units. *J Mol Biol* **272**, 95-105
82. Panchenko AR, Luthey-Schulten Z & Wolynes PG (1996) Foldons, protein structural modules, and exons. *Proc Natl Acad Sci U S A* **93**, 2008-2013
83. Radford SE, Dobson CM & Evans PA (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358**, 302-307
84. Sosnick TR, Mayne L, Hiller R & Englander SW (1994) The barriers in protein folding. *Nat Struct Biol* **1**, 149-156
85. Steiner DF, Cunningham D, Spigelman L & Aten B (1967) Insulin biosynthesis: evidence for a precursor. *Science* **157**, 697-700
86. Sohl JL, Jaswal SS & Agard DA (1998) Unfolded conformations of alpha-lytic protease are more stable than its native state. *Nature* **395**, 817-819
87. Dee DR & Yada RY (2010) The prosegment catalyzes pepsin folding to a kinetically trapped native state. *Biochemistry* **49**, 365-371
88. Wang Z, Mottonen J & Goldsmith EJ (1996) Kinetically controlled folding of the serpin plasminogen activator inhibitor 1. *Biochemistry* **35**, 16443-16448
89. Bryngelson JD, Onuchic JN, Socci ND & Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-195
90. Wagner C & Kiefhaber T (1999) Intermediates can accelerate protein folding. *Proc Natl Acad Sci U S A* **96**, 6716-6721
91. Leopold PE, Montal M & Onuchic JN (1992) Protein folding funnels: a kinetic approach to the sequence-structure relationship. *Proc Natl Acad Sci U S A* **89**, 8721-8725
92. Onuchic JN, Luthey-Schulten Z & Wolynes PG (1997) Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem* **48**, 545-600
93. Onuchic JN, Nymeyer H, Garcia AE, Chahine J & Socci ND (2000) The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Adv Protein Chem* **53**, 87-152
94. Bryngelson JD & Wolynes PG (1989) Intermediates and Barrier Crossing in a Random Energy Model (with Applications to Protein Folding). *J Phys Chem* **93**, 6902-6915
95. Ferreiro DU, Hegler JA, Komives EA & Wolynes PG (2007) Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci U S A* **104**, 19819-19824

96. Dobson CM (2003) Protein folding and misfolding. *Nature* **426**, 884-890
97. Guijarro JI, Sunde M, Jones JA, Campbell ID & Dobson CM (1998) Amyloid fibril formation by an SH3 domain. *Proc Natl Acad Sci U S A* **95**, 4224-4228
98. Onuchic JN & Wolynes PG (2004) Theory of protein folding. *Curr Opin Struct Biol* **14**, 70-75
99. Taketomi H, Ueda Y & Go N (1975) Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effect of specific amino acid sequence represented by specific inter-unit interactions. *Int J Pept Protein Res* **7**, 445-459
100. Dill KA & Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* **4**, 10-19
101. Dyson HJ & Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**, 197-208
102. Gunasekaran K, Tsai CJ, Kumar S, Zanuy D & Nussinov R (2003) Extended disordered proteins: targeting function with less scaffold. *Trends Biochem Sci* **28**, 81-85
103. Uversky VN (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* **11**, 739-756
104. Uversky VN, Gillespie JR & Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **41**, 415-427
105. Netzer WJ & Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388**, 343-349
106. Matthews BW & Remington SJ (1974) The three dimensional structure of the lysozyme from bacteriophage T4. *Proc Natl Acad Sci U S A* **71**, 4178-4182
107. Shank EA, Cecconi C, Dill JW, Marqusee S & Bustamante C (2010) The folding cooperativity of a protein is controlled by its chain topology. *Nature* **465**, 637-640
108. Brandt F, Etchells SA, Ortiz JO, Elcock AH, Hartl FU & Baumeister W (2009) The native 3D organization of bacterial polysomes. *Cell* **136**, 261-271
109. Voss NR, Gerstein M, Steitz TA & Moore PB (2006) The geometry of the ribosomal polypeptide exit tunnel. *J Mol Biol* **360**, 893-906
110. Cabrita LD, Dobson CM & Christodoulou J (2010) Protein folding on the ribosome. *Curr Opin Struct Biol* **20**, 33-45
111. Choi KM, Atkins JF, Gesteland RF & Brimacombe R (1998) Flexibility of the nascent polypeptide chain within the ribosome--contacts from the peptide N-terminus to a specific region of the 30S subunit. *Eur J Biochem* **255**, 409-413
112. Blobel G & Sabatini DD (1970) Controlled proteolysis of nascent polypeptides in rat liver cell fractions. I. Location of the polypeptides within ribosomes. *J Cell Biol* **45**, 130-145
113. Woolhead CA, McCormick PJ & Johnson AE (2004) Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* **116**, 725-736
114. Fulton AB (1982) How crowded is the cytoplasm? *Cell* **30**, 345-347
115. Rivas G, Ferrone F & Herzfeld J (2004) Life in a crowded world. *EMBO Rep* **5**, 23-27
116. van den Berg B, Ellis RJ & Dobson CM (1999) Effects of macromolecular crowding on protein folding and aggregation. *EMBO J* **18**, 6927-6933
117. Monterroso B & Minton AP (2007) Effect of high concentration of inert cosolutes on the refolding of an enzyme: carbonic anhydrase B in sucrose and ficoll 70. *J Biol Chem* **282**, 33452-33458
118. Munishkina LA, Cooper EM, Uversky VN & Fink AL (2004) The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J Mol Recognit* **17**, 456-464
119. Hong J & Gierasch LM (2010) Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J Am Chem Soc* **132**, 10445-10452
120. Ai X, Zhou Z, Bai Y & Choy WY (2006) 15N NMR spin relaxation dispersion study of the molecular crowding effects on protein folding under native conditions. *J Am Chem Soc* **128**, 3916-3917
121. Ladurner AG & Fersht AR (1999) Upper limit of the time scale for diffusion and chain collapse in chymotrypsin inhibitor 2. *Nat Struct Biol* **6**, 28-31
122. Kumar R (2009) Role of naturally occurring osmolytes in protein folding and stability. *Arch Biochem Biophys* **491**, 1-6
123. Bukau B, Weissman J & Horwich A (2006) Molecular chaperones and protein quality control. *Cell* **125**, 443-451

124. Gregersen N, Bross P, Vang S & Christensen JH (2006) Protein misfolding and human disease. *Annu Rev Genomics Hum Genet* **7**, 103-124
125. Hartl FU & Hayer-Hartl M (2009) Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol* **16**, 574-581
126. Hartl FU (1996) Molecular chaperones in cellular protein folding. *Nature* **381**, 571-579
127. Chang HC, Tang YC, Hayer-Hartl M & Hartl FU (2007) SnapShot: molecular chaperones, Part I. *Cell* **128**, 212
128. Tang YC, Chang HC, Hayer-Hartl M & Hartl FU (2007) SnapShot: molecular chaperones, Part II. *Cell* **128**, 412
129. Kramer G, Boehringer D, Ban N & Bukau B (2009) The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat Struct Mol Biol* **16**, 589-597
130. Frydman J, Nimmesgern E, Ohtsuka K & Hartl FU (1994) Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* **370**, 111-117
131. Kerner MJ, Naylor DJ, Ishihama Y, Maier T, Chang HC, Stines AP, Georgopoulos C, Frishman D, Hayer-Hartl M, Mann M, et al. (2005) Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. *Cell* **122**, 209-220
132. Langer T, Lu C, Echols H, Flanagan J, Hayer MK & Hartl FU (1992) Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* **356**, 683-689
133. Wandinger SK, Richter K & Buchner J (2008) The Hsp90 chaperone machinery. *J Biol Chem* **283**, 18473-18477
134. McClellan AJ, Tam S, Kaganovich D & Frydman J (2005) Protein quality control: chaperones culling corrupt conformations. *Nat Cell Biol* **7**, 736-741

Part I – Iron-sulfur clusters and protein stability

2. Iron-Sulfur Clusters as a model of protein conformational stabilizers

2.1. Iron-sulfur clusters: chemistry and function

Despite the extraordinary structural diversity of proteins, amino acids are not able to undergo redox transitions (with the notable exception of tyrosine), catalyze certain reactions and provide enough conformational stabilization in certain proteins. For this reason, proteins incorporate cofactors which render them functional versatility and additional stabilization. Iron-sulfur (FeS) clusters are ubiquitous and biologically essential inorganic cofactors of iron and sulfur thought to have been instrumental prebiotic catalysts and one of the oldest protein cofactors [1]. They usually present di- tri- or tetra-iron nuclearities, though more complex ensembles in which a cuboid cluster is bridged to another component such as nitrogenase's cofactor cluster ([Mo7Fe-9S]) and P-cluster ([8Fe-7S]) are biologically relevant (Figure 2.1) [2]. Protein binding occurs most frequently through cysteine side chains. [2Fe-2S] clusters can also present a mixed coordination where one of the iron atoms is coordinated by two cysteines and the other by two histidine side chains, forming the so-called Rieske cluster [3]. Glutamine, serine or arginine are less common ligands [4].

FeS clusters can be assembled *in vitro* from the spontaneous reaction of iron salts with thiolates and sulfide [5] at millimolar concentrations. However, these concentrations are toxic for cells. The *in vivo* assembly and insertion of FeS clusters in the appropriate apoproteins is a highly regulated and still not completely elucidated process involving desulfurase enzymes for the acquisition of sulfur from cysteine, iron chaperones for the delivery of iron, scaffold proteins for the assembly of the cluster and transfer enzymes (reviewed in [6]). There are several biosynthetic machineries including the NIF system of nitrogen fixing organisms like *Azotobacter vinelandii*, the ISC system of eukaryotic mitochondrial FeS cluster biosynthesis and the bacterial SUF system which operates in bacteria in conditions of iron limitation or oxidative stress, archaea, cyanobacteria and plant chloroplasts.

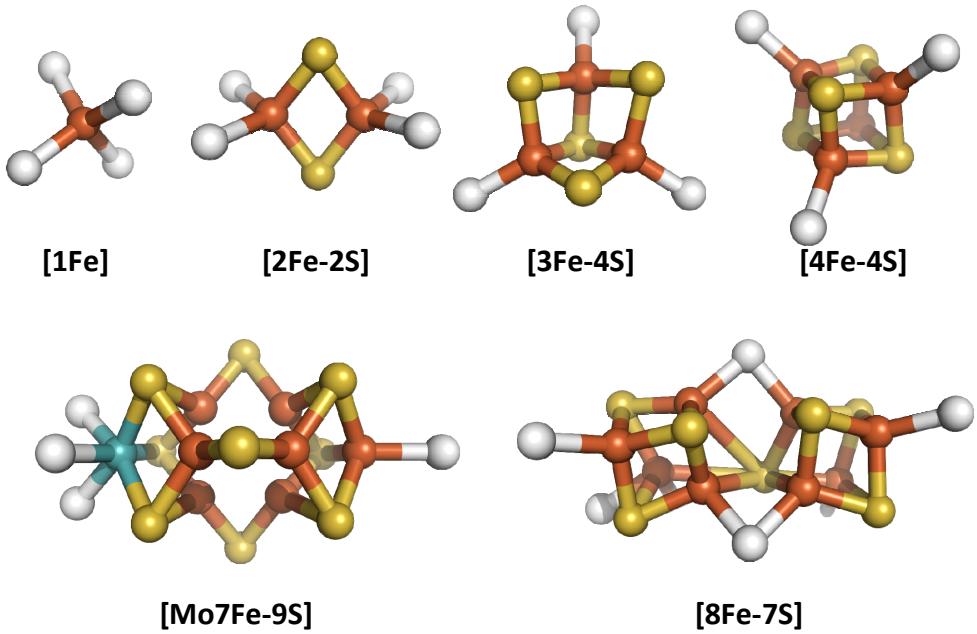


Figure 2.1 – Structure of iron-sulfur clusters. Iron is shown in orange, inorganic sulfur in yellow and molybdenum as blue/green. The protein ligand atoms are white spheres. The figure contains the mononuclear iron in *Desulfovibrio gigas* rubredoxin ([1Fe], 2DSX) and the clusters in *Synechocystis sp.* PCC 6803 ferredoxin ([2Fe-2S], 1OFF), *Acidianus ambivalens* seven-iron ferredoxin ([3Fe-4S] and [4Fe-4S], 2VKR) and *Azotobacter vinelandii* nitrogenase ([Mo7Fe-9S and [8Fe-7S], 1N2C). Clusters are drawn approximately on the same scale.

Eukaryotes express the cytosol-specific CIA machinery, homologous to the mitochondrial ISC system [7].

FeS clusters are highly versatile protein cofactors. They bind to proteins of different folds and sizes, are present in all cellular compartments and are required for a great variety of protein functions [8-10]. The polypeptide chain modulates the intrinsic properties of the clusters, tuning them for the desired functions. This includes the protection of the FeS cluster from its intrinsic oxygen sensitivity to highly stable forms in some proteins [11]. Clusters from different nuclearities can undergo interconversions. Examples of this situation include the conversion of [3Fe-4S] clusters to [4Fe-4S], the oxidative decomposition of [4Fe-4S] to [3Fe-4S] clusters [12-13] and the proposed assembly of [4Fe-4S] from [2Fe-2S] clusters in the FeS clusters biogenesis scaffold proteins.

Table 2.1 – Functions of biological FeS proteins*. From [8].

Function	Proteins	Cluster type
Electron transfer	Ferredoxins, redox enzymes	[2Fe-2S], [3Fe-4S], [4Fe-4S]
Coupled electron/proton transfer	Rieske protein	[2Fe-2S]
	Nitrogenase	[8Fe-7S]
Substrate binding and activation	(de)hydratases	[4Fe-4S]
	Radical SAM enzymes	[4Fe-4S]
	Acetyl-CoA synthase	Ni-Ni-[4Fe-4S], [Ni-4Fe-5S]
	Sulfite reductase	[4Fe-4S]-siroheme
Fe or cluster storage	Ferredoxins	[4Fe-4S]
	Polyferredoxins	[4Fe-4S]
Structural	Endonuclease III	[4Fe-4S]
	MutY	[4Fe-4S]
Regulation of gene expression	SoxR	[2Fe-2S]
	FNR	[4Fe-4S]/[2Fe-2S]
	IRP	[4Fe-4S]
	IscR	[2Fe-2S]
Regulation of enzyme activity	Glutamine PRPP amidotransferase	[4Fe-4S]
	Ferrochelatase	[2Fe-2S]
Disulfide reduction	Ferredoxin:Thioredoxin reductase	[4Fe-4S]
	Heterodisulfide reductase	[4Fe-4S]
Sulfur donor	Biotin synthase	[2Fe-2S]

*SAM, S-adenosylmethionine; Acetyl-CoA, Acetyl coenzymeA; FNR, fumarate and nitrate reduction; IRP, iron-regulatory protein; IscR, iron-sulfur cluster assembly regulatory protein; PRPP, phosphoribosylpyrophosphate.

Most biological roles of FeS clusters are based on the ability of iron to accept or donate electrons, cycling between the ferrous (Fe^{2+}) and ferric (Fe^{3+}) oxidation states. FeS clusters may shuttle between two oxidation states, making them one-electron acceptors/donors. The redox potential of FeS clusters is highly sensitive to the ligand environment in the polypeptide and can be tuned for specific functional requirements where they conduct electrons from one active site to another, like the ones in hydrogenases or oxygenases. The most notable example of electron transfer via FeS clusters occurs in respiratory chain complexes. Complex I features a FeS cluster "wire" which couples NADH oxidation to ubiquinone reduction (Figure 2.2) [14]. Similarly, complex II employs three FeS clusters to transfer electrons resulting from the oxidation of succinate to ubiquinone. Complex III includes a single Rieske-type [2Fe-2S] cluster which transfers electrons from

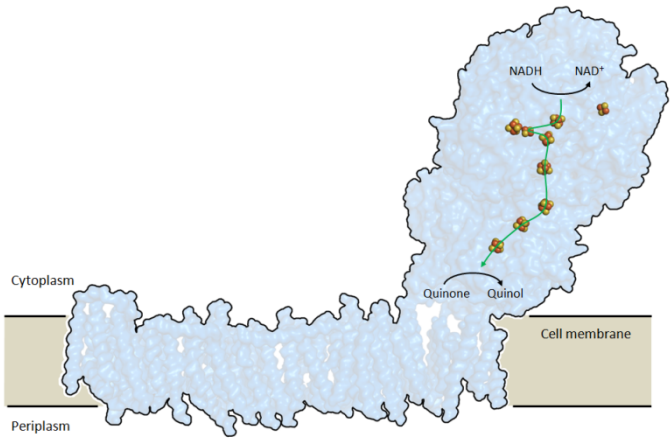


Figure 2.2 – FeS clusters in the electron transfer pathway in respiratory complex I. Structure of the *Thermus thermophilus* complex (3M9S). Seven conserved FeS clusters (one [2Fe-2S] and six [4Fe-4S] clusters) in close proximity act as a "wire" transferring electrons originating from the oxidation of NADH to NAD^+ at the cytoplasm boundary to the quinone oxidation site near the membrane region. The total electron transfer pathway, in green, is around 8 nm. The two FeS clusters not included in the electron transfer pathway are thought to prevent the generation of reactive oxygen species and perform structural roles. Protein is shown as a surface representation; FeS clusters are depicted as spheres. The flavin mononucleotide cofactor at the NADH oxidation site has been omitted for clarity. From [14].

ubiquinone to cytochrome c_1. Electrons coming from the oxidation of fatty acids enter the mitochondrial respiratory chain via the electron transfer flavoprotein:ubiquinone oxidoreductase [15]. Steroid biosynthesis by mitochondrial cytochrome P450 requires electrons from NADPH which are transferred by adrenodoxin, a [2Fe-2S] cluster protein [16]. The exquisite electron transfer properties of FeS clusters are the reason for why they are necessary for enzymes breaking string multiple bonds such as nitrogenase, sulfite and nitrite reductases.

Other chemical reactions where FeS clusters are involved do not involve redox transformations. The first step in the tricarboxylic acid cycle, the major source of reducing equivalents in aerobic respiration, is the conversion of citrate to isocitrate, a step catalyzed by the mitochondrial [4Fe-4S] cluster enzyme aconitase. The cluster is bound by three cysteine sulfurs. The fourth coordination position is occupied by an hydroxyl anion in the enzyme's resting state and is the substrate binding site during the stereo-specific catalysis [17] (Figure 2.3).

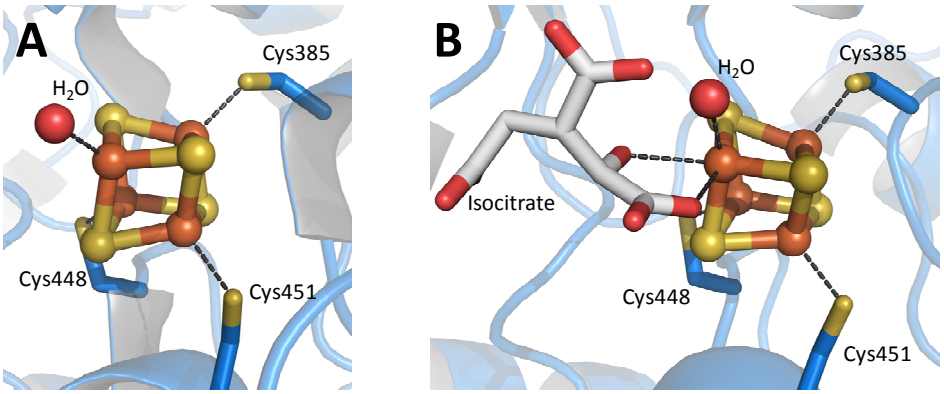


Figure 2.3 – [4Fe-4S] cluster in human aconitase. (A) In the resting state (3B3X), aconitase binds four of the FeS cluster iron atoms (shown as orange spheres). The remainder coordination position is occupied by a water or hydroxide molecule. (B) During catalysis (1AR3) the substrate binds the fourth iron which provides the scaffold for the stereospecific conversion of citrate to isocitrate (shown as sticks).

FeS clusters can also be metabolic precursors. It is believed that the formation of high nuclearity ([3Fe-4S] and [4Fe-4S]) FeS clusters is hierarchic, occurring at the expense of lower nucelarity [2Fe-2S] clusters in the biosynthetic scaffold proteins [8]. Biosynthesis of the molybdenum cofactor (MoCo) present in aldehyde oxidase and xanthine dehydrogenase involves molybdenum cofactor biosynthesis protein 1 (MOCS1), a protein containing a [4Fe-4S] and [3Fe-4S] clusters [18]. It has been suggested that one of the two [4Fe-4S] clusters in lipoate synthase serves as the source of sulfur in the formation of lipoic acid [19]. Nucleic acid metabolism is highly dependent on FeS enzymes. Humans have three FeS enzymes in this pathway: dihydropyrimidine dehydrogenase (DPYD), phosphoribosylpyrophosphate amidotransferase (PPAT) and xanthine dehydrogenase (XDH) [10]. Curiously, the sole [4Fe-4S] cluster in PPAT does not seem to be required for activity but rather for amino-terminal processing and stabilization [20-21]. FeS proteins are also involved in DNA and RNA synthesis and processing. ATP binding cassette protein E1 binds two [4Fe-4S] clusters and is involved in protein translation initiation, rRNA processing, ribosomal assembly [22]. The BACH1 helicase, the PRIM2 DNA primase [23],

the ERCC2 helicase, the NTHL1 DNA endonuclease [24] and the MUTYH DNA glycosidase are all [4Fe-4S] proteins. At the level genetic regulation FeS clusters also play a role. The *E. coli* FNR transcription factor is responsible for the transcription of genes involved in anaerobic energy transduction systems which senses the oxygen concentration through its [4Fe-4S] cluster which converts to [2Fe-2S] in the presence of oxygen, leading to the dissociation of the active FNR dimer and gene transcription inhibition [25].

Iron metabolism requires the action of several FeS proteins: ferrochelatase – the iron chaperone which inserts Fe^{2+} in protoporphyrin IX to yield heme – has a [2Fe-2S] cluster. The synthesis of the major iron metabolism proteins – ferritin (storage) and transferrin (transport) – is controlled by the amount of iron in the cell which is sensed by the cytosolic aconitase, a bacteria and eukaryotic bifunctional protein containing a [4Fe-4S] cluster like its mitochondrial homolog. In high-iron conditions, the protein is free in the cytosol and ferritin is synthesized by the cell; transferrin mRNA is unstable and is rapidly degraded. Under conditions of low iron, the [4Fe-4S] cluster dissociates from the protein which then binds Iron-responsive elements (IREs) in mRNA. This represses ferritin mRNA translation and stabilizes transferrin mRNA, which drives protein synthesis.

2.2. Iron-sulfur, protein folds and structural plasticity

FeS clusters are harbored in nearly 50 unique protein folds [26]. Over 90% of these folds harbor low-potential [2Fe-2S] or [4Fe-4S] clusters. FeS clusters are highly versatile in the sense that the interaction between the polypeptide and the cluster determines which kind of cluster binds to the protein and the polypeptide determines the redox potential and reactivity of the cluster. Additionally, the biological importance of FeS clusters is based on their ability to cycle between redox states, coordination configurations and nuclearity and stabilize the native conformation of proteins.

2.2.1. Rubredoxin

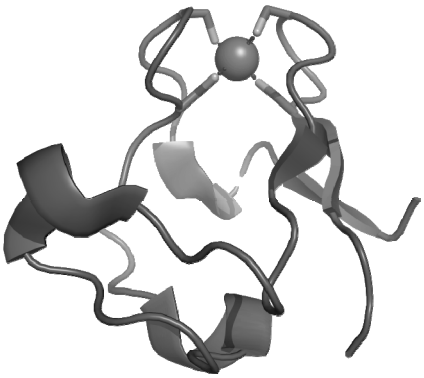


Figure 2.4 - Rubredoxin fold. *Desulfovibrio gigas* rubredoxin (2DSX). The iron atom is shown as a sphere and the cysteine ligands as sticks.

Rubredoxin is a ~50 amino acid residue bacterial and archaeal protein containing a single tetrahedrally-coordinated mononuclear iron (Figure 2.4). It is involved in electron transfer, sometimes replacing ferredoxin [27]. The highly conserved fold belongs to the α+β family, with 2 α-helices and 2-3 β-strands. The rubredoxin fold occurs in multidomain proteins like desulforedoxin [28], ruberythrin [29] and superoxide reductase [30] and in the FeS cluster domain of Rieske proteins [31]. In the latter case, the iron binding site accommodates the Rieske [2Fe-2S] cluster without fold distortion. A rubredoxin iron site can be converted into a [2Fe-2S] site and vice versa by specific point mutations [32-33].

2.2.2. [2Fe-2S] plant- and vertebrate-type ferredoxin

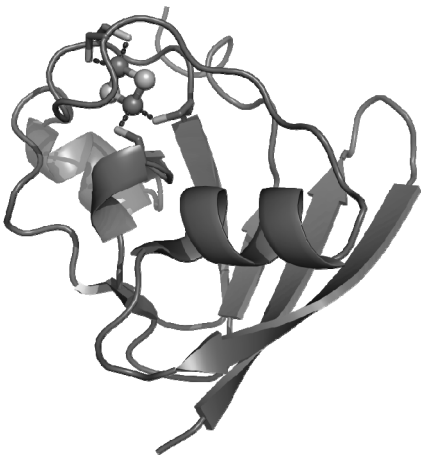


Figure 2.5 - [2Fe-2S] ferredoxin. From *Anabaena* PCC 7119 (1CZP).

These proteins constitute a functionally diverse group. Plant-type ferredoxins carry electrons resulting from water oxidation at photosystem I to several enzymes. The bacterial (*e. g.* putidaredoxin) and vertebrate (*e. g.* adrenodoxin) ferredoxins donate electrons to hydroxylating enzymes like cytochrome P450. The ISC FeS biosynthesis machinery includes a ferredoxin electron transporter of this

type. This type of ferredoxins has a low potential cluster (-150 to -450 mV).

These proteins are small (~100 residues) belonging to the α+β fold class. The [2Fe-2S] cluster is located close to the surface in a loop region. The opposite side of the protein forms a ubiquitin-like motif with a four stranded β-sheet and helical segments (Figure 2.5). This type of ferredoxin occurs as a domain in larger dehydrogenase, reductase and oxidase enzymes. The apically located [2Fe-2S] cluster is not essential for protein stability. The stabilizing contributions for the polypeptide and the cluster itself are mainly based in the opposite, secondary structure-rich part of the protein [34].

2.2.3. Rieske [2Fe-2S] ferredoxin

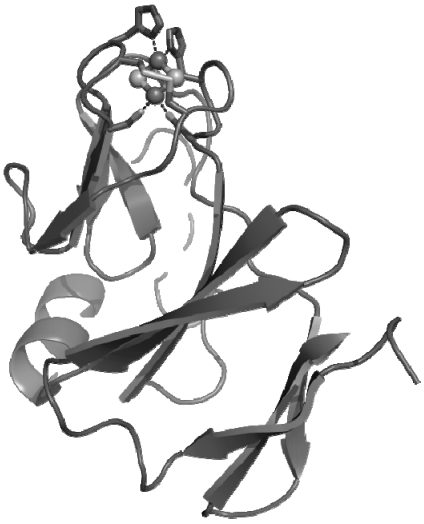


Figure 2.6 – Rieske fold. Soluble domain of the bovine mitochondrial Rieske protein in cytochrome bc_1 (1RIE). The disulfide bridge connecting the two cluster binding loops is depicted as sticks. The membrane anchor is not included in the structure.

Rieske proteins contain a [2Fe-2S] cluster where one of the iron atoms is coordinated by two cysteine side chains and the other by two histidine side chains. They are obligatory components of the membrane-bound respiratory (cytochrome bc_1) and photosynthetic (cytochrome b_6f) electron transfer complexes. Soluble Rieske proteins, named Rieske ferredoxins, are components of oxygenase enzymatic systems or electron carriers.

Despite having a cluster with the same nuclearity as [2Fe-2S] plant- and vertebrate-type ferredoxins, Rieske proteins have a distinct fold and specific functions. The Rieske fold consists of three stacked β-sheets, of which the upper one contains the cluster ligands in two loops, each one containing a cysteine and a histidine residue. In respiratory and photosynthetic Rieske proteins the cluster binding loops are

further cross-linked by a disulfide bridge (Figure 2.6) which is crucial for protein stability and activity, since mutations in the intervening cysteines precludes cluster formation. The cluster histidine ligands are solvent exposed and are responsible for the highly pH-dependent redox potential of the Rieske cluster. Membrane-bound Rieske proteins have an N-terminal hydrophobic extension which is inserted in the membrane. Soluble Rieske proteins have the same fold as the soluble domain of the membrane-bound counterparts with the exception of having an helical segment removed. The fold around the cluster is highly similar to the rubredoxin one. The unique [2Fe-2S] coordination implies upshifted redox potentials (+100 to +400 mV for respiratory and photosynthetic proteins and around -100 mV for dioxygenase components).

2.2.4. Thioredoxin-like [2Fe-2S] ferredoxin

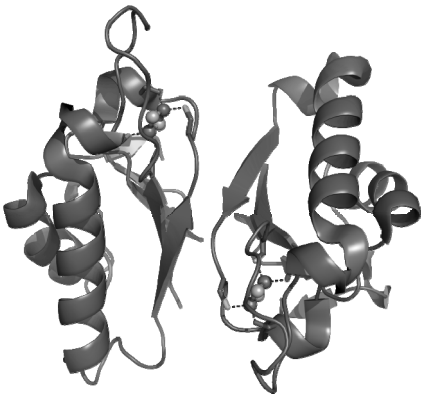


Figure 2.7 – Thioredoxin-like [2Fe-2S] fold. From *Aquifex aeolicus* (11M2D). Dimeric structure. The FeS cluster is shown in ball and sticks. The ligands are shown as sticks.

Thioredoxin-like ferredoxins are only found in bacteria as components of hydrogenase and NADH:ubiquinone oxidoreductase [35] and have been proposed to be involved in nitrogen metabolism [36]. Unlike thioredoxin, the β-sheet is wrapped by α-helices on one side (Figure 2.7). The other side of the sheet is involved in homodimerization in the *Aquifex aeolicus* protein. The redox potential of the [2Fe-2S] cluster is around -300 mV.

2.2.5. High-potential [4Fe-4S] protein

[4Fe-4S] high-potential iron proteins (HiPIP) are mostly found in photosynthetic bacteria [37] where they donate electrons to the tetraheme

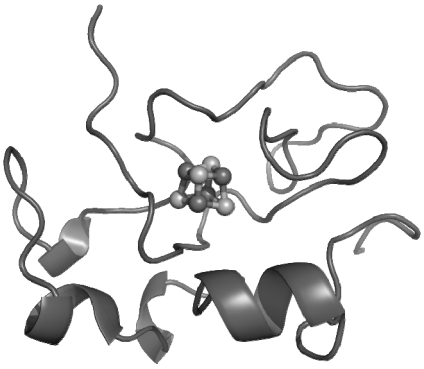


Figure 2.8 – High Potential Iron-sulfur protein (HiPIP). From *Allochromatium vinosum* (1B0Y).

cytochrome [38]. Their name is derived from the [4Fe-4S] cluster reduction potential (+100 to +400 mV) [26], resulting from the incorporation of the cluster in the protein's hydrophobic core. The most distinctive structural feature of these proteins is the little secondary structure content in their globular conformation (Figure 2.8).

2.2.6. Di-cluster [4Fe-4S] ferredoxin

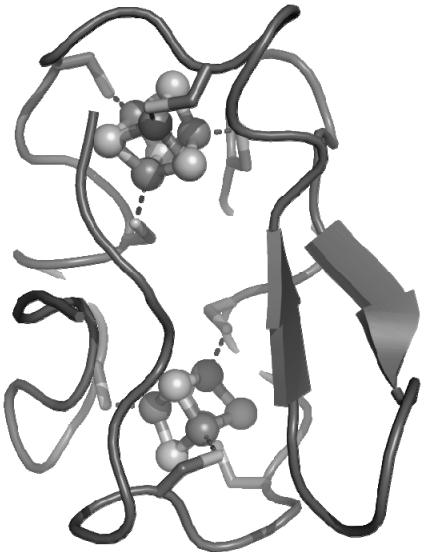


Figure 2.9 - [4Fe-4S] dicluster ferredoxin. From *Clostridium pasteurianum* (1CLF).

These proteins are involved in anaerobic metabolic pathways and in the most reducing steps of photosynthetic and aerobic electron transport chains due to the low reduction potential of the [4Fe-4S] clusters (distinct for each one, ranging from -150 to -700 mV) [26]. The reference for these proteins is *Clostridium pasteurianum* ferredoxin. The [4Fe-4S] clusters contribute significantly to protein stability because of the tight wrapping of the polypeptide around them. The two clusters make up the protein core in a small protein (55 amino acid residues) and this precludes the necessity of additional stabilizing forces. The structure has a pseudo two-fold symmetry, with each half harboring a cluster. This protein fold is the only one harboring a pair of clusters rather than just one.

The ferredoxin fold has been proposed to be one of the most primitive ones due to its simple amino acid composition [39]. The high stability of some fold representatives (*e. g.* [40]) also agrees with the thermophilic origin of life.

The characteristic topology for these proteins, $(\beta\alpha\beta)_2$, is referred to as the ferredoxin fold (Figure 2.9). This fold constitutes a core structure from where several modular variations exist. These include the C-terminal extensions in *Pseudomonas aeruginosa* [41] and *Azotobacter vinelandii* [42], the N-terminal extensions of *Sulfolobales* including an additional Zn^{2+} site [43-44] or a stabilizing hydrophobic core [45] or bacterial ferredoxins in which one of the clusters has been lost [42, 46-47].

2.3. Folding and stability of iron-sulfur proteins

One must analyze two aspects when considering the conformational interplay between FeS clusters and proteins. One the one hand, FeS clusters are intrinsically oxygen sensitive [48] and need an appropriate ligand environment to be stabilized; on the other hand, the incorporation of the cluster is frequently a long range polypeptide cross-linker which is essential for protein stability. FeS clusters are usually embedded in a single protein domain, which provides the relatively rigid and stabilizing ligand environment. Solvent exposure of FeS clusters emerges from functional requirements, like in soluble electron carriers or in sensor or regulatory proteins.

Despite the significant variety of folds able to accommodate FeS clusters, each fold usually cannot bind different clusters because of strict steric restrictions. The exceptions are the superimposable rubredoxin and Rieske folds around the [2Fe-2S] cluster. Nevertheless, cluster swapping only occurs under mutagenic perturbations. Nitrogenase contains a native [4Fe-4S] cluster [2] which can be cleaved into [2Fe-2S] clusters but as a result of protein denaturation [49] or by glycerol-induced decomposition in the Mg^{2+}-bound state [50].

Apart from their functional role, FeS clusters are key structural elements stabilizing the native protein conformation [40, 51]. Their multi-ligand coordination establishes long range cross-links in the polypeptide, restricts conformational entropy, favors folding and stabilizes the native state. Many FeS proteins are not able to acquire the native or native-like fold in the absence of the cognate FeS clusters [52] or require the conversion between different cluster nuclearities for functional roles [25].

2.4. Folding and stability of *Acidianus ambivalens* FeS proteins

One of the most intriguing open topics in the protein folding field currently is the understanding of the molecular determinants of protein stability. That is, what are the amino acid sequence features and the interactions involved in selecting and maintaining the native conformation. In the case of cofactor-binding proteins, this issue is enlarged by considering the effect of the cofactor and polypeptide-cofactor interactions. Three main strategies have been adopted to tackle this issue:

1. Comparative analysis of the folding properties of wild type and single point amino acid mutants;
2. Study of the folding properties of single protein domains or small peptides derived from actual proteins;
3. Study of the folding properties of intrinsically highly stable proteins.

Under the framework of the third strategy, proteins originating from extremophiles – organisms thriving at extreme temperature, pH, salinity or pressure conditions – are obvious research models because it is assumed that the organism's survivability is set by the least stable of its components. The majority of extremophilic organisms are Archaea, which constitute the deepest branch in Woese's tree of life [53-54]. The adaptation to extreme conditions together with the primordial origin of these organisms has been linked to the origin of life on Earth. In this sense, archaeal FeS proteins are

unique models for addressing the molecular determinants of protein stabilization because of their primordial origin and features (*e. g.* the $(\beta\alpha\beta)_2$ ferredoxin fold is considered to be one of the oldest folds), spectroscopically active cofactors and small size of the FeS binging domains. Some organisms such as thermoacidophilic *Sulfolobales* family members intrinsically express such high amounts of FeS proteins that they can be spectroscopically detected in intact cells [55]. This has allowed the unequivocal identification of [3Fe-4S][4Fe-4S] ferredoxins *in vivo* and the establishment that [3Fe-4S] clusters are physiologically relevant and not decomposed [4Fe-4S] clusters.

Two members of the *Sulfolobales* family have been extensively studied in terms of the folding properties of their FeS proteins, namely ferredoxins: *Acidianus ambivalens* and *Sulfolobus metallicus*. These ferredoxins have a fold identical to that of the canonical dicluster [4Fe-4S] ferredoxin, but one of the clusters is replaced for a [3Fe-4S] one (Figure 2.10). The family of seven-iron ferredoxins is intrinsically very stable. For the *Thermus thermophilus* homolog T_m = 114°C and C_m = 6.5 M GuHCl at pH 7 [56]. Ionic interactions were suggested to be a major stabilization due to substantial decreased stability at lower pH [56-57]. Despite their small size, seven-iron ferredoxins unfold irreversibly. Two mechanisms are thought to contribute to this outcome: first, the clusters constitute a considerable part of the protein hydrophobic core, which is densely packed in thermostable proteins; second, the clusters should decompose after release from the polypeptide, precluding reversibility.

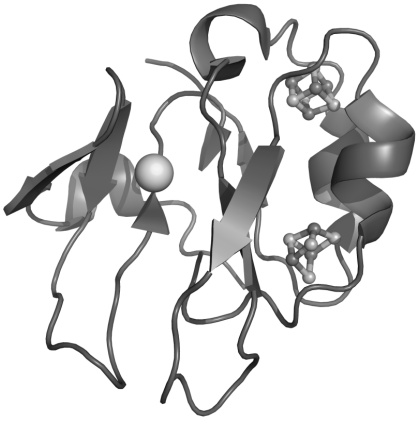


Figure 2.10 – Seven iron ferredoxin from *Acidianus ambivalens*. The structure includes the [3Fe-4S] and [4Fe-4S] clusters and the Zn^{2+} site. PDB: 3VKR.

Acidianus ambivalens is a thermoacidophile growing optimally at

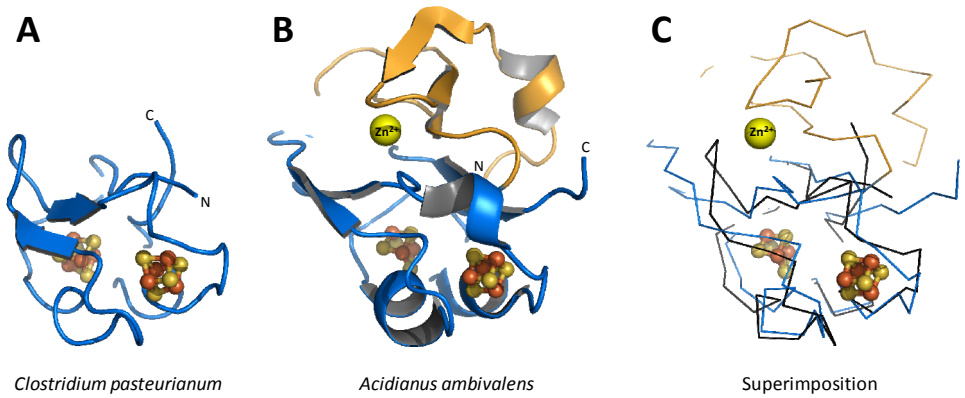


Figure 2.11 – Unique features of the *Acidianus ambivalens* [3Fe-4S][4Fe-4S] ferredoxin. (A) dicluster [4Fe-4S] ferredoxin from *Clostridium pasteurianum* (1CLF). (B) [3Fe-4S][4Fe-4S] ferredoxin from *Acidianus ambivalens* (2VKR). The N-terminal extension is shown in orange and the Zn^{2+} ion as a yellow sphere. (C) Backbone superimposition of the two structures. The *C. pasteurianum* structure is in black and the *A. ambivalens* is in blue/orange. Only the FeS clusters of the *A. ambivalens* protein are shown. The clusters from the *C. pasteurianum* protein are superimposable.

80 °C and pH 2.5 [58]. It expresses a [3Fe-4S][4Fe-4S] ferredoxin whose fold is identical to that of the canonical dicluster [4Fe-4S] ferredoxin, except for the presence of an N-terminal extension harboring a His/Asp Zn^{2+} site (Figure 2.11).

The folding properties of this protein have been extensively studied and a comprehensive view of the unfolding process has been achieved. The relative contribution of the polypeptide versus the clusters has been analyzed taking advantage of the distinct spectroscopic properties of both. At pH 7 the protein is highly stable and remains folded even in 8 M GuHCl [59]. The T_m is ~120°C in the 5.5 to 8 pH range [57]. Changing pH to the protein's isoelectric point (pH 3.5), in which the electrostatic contributions for protein stabilization are minimized, the protein undergoes a conformational change which does not affect the FeS clusters. The latter then unfold at a higher temperature than the polypeptide (T_m = 72°C versus 52°C). Below pH 2 the FeS clusters disintegrate spontaneously [57]. Overall, these observations

show that the FeS clusters in this ferredoxin are essential for the maintenance of the native conformation.

Thermal denaturation of the ferredoxin at pH 2.5 yields a FeS- and Zn^{2+}-less stable molten globule conformation which unfolds reversibly and is only marginally less stable than the holo state ($\Delta\Delta G$ ~1.5 kcal/mol) [60]. Much evidence exists in the literature suggesting that molten globule states may be ubiquitous folding intermediates [61] and that FeS protein molten globule like conformations are good candidates for being the acceptors of FeS clusters from the biosynthetic machinery. The association of conformational plasticity with the mechanism of FeS clusters insertion into apoproteins is further highlighted in IscU, the scaffold protein from where pre-formed FeS clusters are transferred to the target apoproteins. It hosts a [2Fe-2S] cluster [62] and adopts a dynamic molten globule-like state which is relevant in the context of protein-protein interactions [63]. Accordingly, and together with the stability of this state, the molten globule state was surmised to be the likely conformational state where the FeS clusters are incorporated in the apoprotein [60].

Sulfolobus metallicus was found to express identical amounts of two [3Fe-4S][4Fe-4S] ferredoxins – an observation repeated in *A. ambivalens* [64] – sharing most biochemical and spectroscopic features [11] but differing in the *Acidianus ambivalens*-like N-terminal extension (Figure 2.12). Ferredoxin A (FdA) includes the His/Asp Zn^{2+} site also found in the *A. ambivalens* protein. The ferredoxin B isoforms (FdB) differs in this region by several polar-to-hydrophobic and charged-to-hydrophobic substitutions which create a compact hydrophobic core. The kinetic stability of the two proteins is similar, with no changes after 72 h incubation at 70°C [11]. Surprisingly, the Zn^{2+}-binding isoforms was found to be less temperature stable (105.9°C *versus* 113.1°C at pH 7), revealing that an evolution-selected hydrophobic contribution can be more stabilizing than the extrinsic Zn^{2+} cross-linking

[45]. This is however uncommon in *Archaea*, as in *A. ambivalens* the Zn^{2+}-containing isoform is less stable than the Zn^{2+}-lacking one (108°C versus 116°C, respectively, at pH 7) [57, 64].

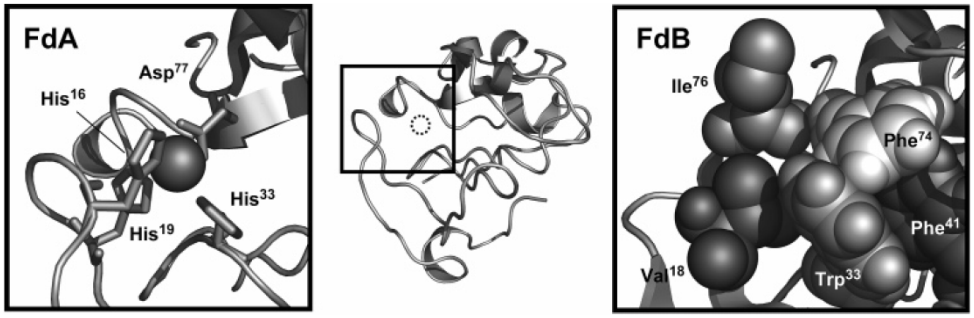


Figure 2.12 – N-terminal extension in *Sulfolobus metallicus* ferredoxin isoforms. The central figure shows the common fold of FdA and FdB. The square-delimited region is blown up in the lateral panels. In FdA the region comprises the conserved His/Asp Zn^{2+} ligands. In FdB this region is filled with tightly packed bulky hydrophobic side chains. From [45].

2.5. References

1. Wachtershauser G (1988) Before enzymes and templates: theory of surface metabolism. *Microbiol Rev* **52**, 452-484
2. Rees DC (2002) Great metalloclusters in enzymology. *Annu Rev Biochem* **71**, 221-246
3. Rieske JS, Maclennan DH & Coleman R (1964) Isolation and properties of an iron-protein from the (reduced coenzyme Q)-cytochrome C reductase complex of the respiratory chain. *Biochem Biophys Res Commun* **15**, 338–344
4. Moulis J-M, Davasse V, Golinelli M-P, Meyer J & Quinkal I (1996) The coordination sphere of iron-sulfur clusters: lessons from site-directed mutagenesis experiments. *J Biol Inorg Chem* **1**, 2-14
5. Herskovitz T, Averill BA, Holm RH, Ibers JA, Phillips WD & Weiher JF (1972) Structure and properties of a synthetic analogue of bacterial iron--sulfur proteins. *Proc Natl Acad Sci U S A* **69**, 2437-2441
6. Bandyopadhyay S, Chandramouli K & Johnson MK (2008) Iron-sulfur cluster biosynthesis. *Biochem Soc Trans* **36**, 1112-1119
7. Lill R & Muhlenhoff U (2008) Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annu Rev Biochem* **77**, 669-700
8. Johnson DC, Dean DR, Smith AD & Johnson MK (2005) Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem* **74**, 247-281
9. Johnson MK (1998) Iron-sulfur proteins: new roles for old clusters. *Current opinion in chemical biology* **2**, 173-181
10. Sheftel A, Stehling O & Lill R (2010) Iron-sulfur proteins in health and disease. *Trends Endocrinol Metab* **21**, 302-314
11. Gomes CM, Faria A, Carita JC, Mendes J, Regalla M, Chicau P, Huber H, Stetter KO & Teixeira M (1998) Di-cluster, seven-iron ferredoxins from hyperthermophilic Sulfolobales. *J Biol Inorg Chem* **3**, 499-507
12. Bell SH, Dickson DP, Johnson CE, Cammack R, Hall DO & Rao KK (1982) Mossbauer spectroscopic evidence for the conversion of [4 Fe--4 S] clusters in Bacillus stearothermophilus ferredoxin into [3 Fe--3 S] clusters. *FEBS Lett* **142**, 143-146
13. Teixeira M, Moura I, Xavier AV, Moura JJ, LeGall J, DerVartanian DV, Peck HD, Jr. & Huynh BH (1989) Redox intermediates of Desulfovibrio gigas [NiFe] hydrogenase generated under hydrogen. Mossbauer and EPR characterization of the metal centers. *J Biol Chem* **264**, 16435-16450

14. Efremov RG, Baradaran R & Sazanov LA (2010) The architecture of respiratory complex I. *Nature* **465**, 441-445
15. Ruzicka FJ & Beinert H (1975) A new membrane iron-sulfur flavoprotein of the mitochondrial electron transfer system. The entrance point of the fatty acyl dehydrogenation pathway? *Biochem Biophys Res Commun* **66**, 622-631
16. Ewen KM, Kleser M & Bernhardt R (2010) Adrenodoxin: The archetype of vertebrate-type [2Fe-2S] cluster ferredoxins. *Biochim Biophys Acta*
17. Beinert H, Kennedy MC & Stout CD (1996) Aconitase as Ironminus signSulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chem Rev* **96**, 2335-2374
18. Hanzelmann P, Hernandez HL, Menzel C, Garcia-Serres R, Huynh BH, Johnson MK, Mendel RR & Schindelin H (2004) Characterization of MOCS1A, an oxygen-sensitive iron-sulfur protein involved in human molybdenum cofactor biosynthesis. *J Biol Chem* **279**, 34721-34732
19. Jarrett JT (2005) The novel structure and chemistry of iron-sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase. *Arch Biochem Biophys* **433**, 312-321
20. Makaroff CA, Paluh JL & Zalkin H (1986) Mutagenesis of ligands to the [4 Fe-4S] center of Bacillus subtilis glutamine phosphoribosylpyrophosphate amidotransferase. *J Biol Chem* **261**, 11416-11423
21. Grandoni JA, Switzer RL, Makaroff CA & Zalkin H (1989) Evidence that the iron-sulfur cluster of Bacillus subtilis glutamine phosphoribosylpyrophosphate amidotransferase determines stability of the enzyme to degradation in vivo. *J Biol Chem* **264**, 6058-6064
22. Kispal G, Sipos K, Lange H, Fekete Z, Bedekovics T, Janaky T, Bassler J, Aguilar Netz DJ, Balk J, Rotte C, et al. (2005) Biogenesis of cytosolic ribosomes requires the essential iron-sulphur protein Rli1p and mitochondria. *EMBO J* **24**, 589-598
23. Weiner BE, Huang H, Dattilo BM, Nilges MJ, Fanning E & Chazin WJ (2007) An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. *J Biol Chem* **282**, 33444-33451
24. Ikeda S, Biswas T, Roy R, Izumi T, Boldogh I, Kurosky A, Sarker AH, Seki S & Mitra S (1998) Purification and characterization of human NTH1, a homolog of Escherichia coli endonuclease III. Direct identification of Lys-212 as the active nucleophilic residue. *J Biol Chem* **273**, 21585-21593
25. Spiro S & Guest JR (1990) FNR and its role in oxygen-regulated gene expression in Escherichia coli. *FEMS Microbiol Rev* **6**, 399-428
26. Meyer J (2008) Iron-sulfur protein folds, iron-sulfur chemistry, and evolution. *J Biol Inorg Chem* **13**, 157-170
27. Lee WY, Brune DC, LoBrutto R & Blankenship RE (1995) Isolation, characterization, and primary structure of rubredoxin from the photosynthetic bacterium, Heliobacillus mobilis. *Arch Biochem Biophys* **318**, 80-88
28. Archer M, Huber R, Tavares P, Moura I, Moura JJ, Carrondo MA, Sieker LC, LeGall J & Romao MJ (1995) Crystal structure of desulforedoxin from Desulfovibrio gigas determined at 1.8 A resolution: a novel non-heme iron protein structure. *J Mol Biol* **251**, 690-702
29. deMare F, Kurtz DM, Jr. & Nordlund P (1996) The structure of Desulfovibrio vulgaris rubrerythrin reveals a unique combination of rubredoxin-like FeS4 and ferritin-like diiron domains. *Nat Struct Biol* **3**, 539-546
30. Yeh AP, Hu Y, Jenney FE, Jr., Adams MW & Rees DC (2000) Structures of the superoxide reductase from Pyrococcus furiosus in the oxidized and reduced states. *Biochemistry* **39**, 2499-2508
31. Link TA (2008) Fe-S Rieske center. In *Handbook of metalloproteins* (Messerschmidt A, Huber R, Poulos T & Wieghardt K, eds). John Wiley & Sons.
32. Meyer J, Gagnon J, Gaillard J, Lutz M, Achim C, Munck E, Petillot Y, Colangelo CM & Scott RA (1997) Assembly of a [2Fe-2S]2+ cluster in a molecular variant of Clostridium pasteurianum rubredoxin. *Biochemistry* **36**, 13374-13380
33. Iwasaki T, Kounosu A, Tao Y, Li Z, Shokes JE, Cosper NJ, Imai T, Urushiyama A & Scott RA (2005) Rational design of a mononuclear metal site into the archaeal Rieske-type protein scaffold. *J Biol Chem* **280**, 9129-9134
34. Morales R, Charon MH, Hudry-Clergeon G, Petillot Y, Norager S, Medina M & Frey M (1999) Refined X-ray structures of the oxidized, at 1.3 A, and reduced, at 1.17 A, [2Fe-2S] ferredoxin from the cyanobacterium Anabaena PCC7119 show redox-linked conformational changes. *Biochemistry* **38**, 15764-15773
35. Vignais PM, Billoud B & Meyer J (2001) Classification and phylogeny of hydrogenases. *FEMS Microbiol Rev* **25**, 455-501

36. Meyer J (2001) Ferredoxins of the third kind. *FEBS Lett* **509**, 1-5
37. Bartsch RG (1978) Purification of (4Fe-4S)1--2--ferredoxins (high-potential iron--sulfur proteins) from bacteria. *Methods Enzymol* **53**, 329-340
38. Ciurli S & Musiani F (2005) High potential iron-sulfur proteins and their role as soluble electron carriers in bacterial photosynthesis: tale of a discovery. *Photosynth Res* **85**, 115-131
39. Eck RV & Dayhoff MO (1966) Evolution of the structure of ferredoxin based on living relics of primitive amino Acid sequences. *Science* **152**, 363-366
40. Leal SS, Teixeira M & Gomes CM (2004) Studies on the degradation pathway of iron-sulfur centers during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability. *J Biol Inorg Chem* **9**, 987-996
41. Giastas P, Pinotsis N, Efthymiou G, Wilmanns M, Kyritsis P, Moulis JM & Mavridis IM (2006) The structure of the 2[4Fe-4S] ferredoxin from Pseudomonas aeruginosa at 1.32-A resolution: comparison with other high-resolution structures of ferredoxins and contributing structural features to reduction potential values. *J Biol Inorg Chem* **11**, 445-458
42. Schipke CG, Goodin DB, McRee DE & Stout CD (1999) Oxidized and reduced Azotobacter vinelandii ferredoxin I at 1.4 A resolution: conformational change of surface residues without significant change in the [3Fe-4S]+/0 cluster. *Biochemistry* **38**, 8228-8239
43. Fujii T, Hata Y, Wakagi T, Tanaka N & Oshima T (1996) Novel zinc-binding centre in thermoacidophilic archaeal ferredoxins. *Nat Struct Biol* **3**, 834-837
44. Frazao C, Aragao D, Coelho R, Leal SS, Gomes CM, Teixeira M & Carrondo MA (2008) Crystallographic analysis of the intact metal centres [3Fe-4S](1+/0) and [4Fe-4S](2+/1+) in a Zn(2+) -containing ferredoxin. *FEBS Lett* **582**, 763-767
45. Rocha R, Leal SS, Teixeira VH, Regalla M, Huber H, Baptista AM, Soares CM & Gomes CM (2006) Natural domain design: enhanced thermal stability of a zinc-lacking ferredoxin isoform shows that a hydrophobic core efficiently replaces the structural metal site. *Biochemistry* **45**, 10376-10384
46. Kissinger CR, Sieker LC, Adman ET & Jensen LH (1991) Refined crystal structure of ferredoxin II from Desulfovibrio gigas at 1.7 A. *J Mol Biol* **219**, 693-715
47. Macedo-Ribeiro S, Darimont B, Sterner R & Huber R (1996) Small structural changes account for the high thermostability of 1[4Fe-4S] ferredoxin from the hyperthermophilic bacterium Thermotoga maritima. *Structure* **4**, 1291-1301
48. Venkateswara Rao P & Holm RH (2004) Synthetic analogues of the active sites of iron-sulfur proteins. *Chem Rev* **104**, 527-559
49. Anderson GL & Howard JB (1984) Reactions with the oxidized iron protein of Azotobacter vinelandii nitrogenase: formation of a 2Fe center. *Biochemistry* **23**, 2118-2122
50. Sen S, Igarashi R, Smith A, Johnson MK, Seefeldt LC & Peters JW (2004) A conformational mimic of the MgATP-bound "on state" of the nitrogenase iron protein. *Biochemistry* **43**, 1787-1797
51. Dilg AW, Grantner K, Iakovleva O, Parak FG, Babini E, Bertini I, Capozzi F, Luchinat C & Meyer-Klaucke W (2002) Dynamics of wild-type HiPIPs: a Cys77Ser mutant and a partially unfolded HiPIP. *J Biol Inorg Chem* **7**, 691-703
52. Pilon M, de Kruijff B & Weisbeek PJ (1992) New insights into the import mechanism of the ferredoxin precursor into chloroplasts. *J Biol Chem* **267**, 2548-2556
53. Woese CR & Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* **74**, 5088-5090
54. Woese CR, Magrum LJ & Fox GE (1978) Archaebacteria. *J Mol Evol* **11**, 245-251
55. Teixeira M, Batista R, Campos AP, Gomes C, Mendes J, Pacheco I, Anemuller S & Hagen WR (1995) A seven-iron ferredoxin from the thermoacidophilic archaeon Desulfurolobus ambivalens. *Eur J Biochem* **227**, 322-327
56. Griffin S, Higgins CL, Soulimane T & Wittung-Stafshede P (2003) High thermal and chemical stability of Thermus thermophilus seven-iron ferredoxin. Linear clusters form at high pH on polypeptide unfolding. *Eur J Biochem* **270**, 4736-4743
57. Leal SS & Gomes CM (2008) On the relative contribution of ionic interactions over iron-sulfur clusters to ferredoxin stability. *Biochim Biophys Acta* **1784**, 1596-1600
58. Zillig W, Yeats S, Holz I, Bock A, Gropp F, Rettenberger M & Lutz S (1985) Plasmid-related anaerobic autotrophy of the novel archaebacterium Sulfolobus ambivalens. *Nature* **313**, 789-791
59. Wittung-Stafshede P, Gomes CM & Teixeira M (2000) Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens. *J Inorg Biochem* **78**, 35-41

60. Leal SS & Gomes CM (2007) Studies of the molten globule state of ferredoxin: structural characterization and implications on protein folding and iron-sulfur center assembly. *Proteins* **68**, 606-616
61. Ptitsyn OB, Pain RH, Semisotnov GV, Zerovnik E & Razgulyaev OI (1990) Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett* **262**, 20-24
62. Mansy SS, Wu G, Surerus KK & Cowan JA (2002) Iron-sulfur cluster biosynthesis. Thermatoga maritima IscU is a structured iron-sulfur cluster assembly protein. *J Biol Chem* **277**, 21397-21404
63. Bertini I, Cowan JA, Del Bianco C, Luchinat C & Mansy SS (2003) Thermotoga maritima IscU. Structural characterization and dynamics of a new class of metallochaperone. *J Mol Biol* **331**, 907-924
64. Janssen S, Trincao J, Teixeira M, Schafer G & Anemuller S (2001) Ferredoxins from the archaeon Acidianus ambivalens: overexpression and characterization of the non-zinc-containing ferredoxin FdB. *Biol Chem* **382**, 1501-1507

3. Role of a novel disulfide bridge within the all-beta fold of soluble Rieske proteins

This chapter was published in

Botelho HM, Leal SS, Veith A, Prosinecki V, Bauer C, Fröhlich R, Kletzin A, Gomes CM (2010) Role of a novel disulfide bridge within the all-beta fold of soluble Rieske proteins. *J Biol Inorg Chem* 15, 271-281

Cloning and expression were performed by Andreas Veith, Christian Bauer, Renate Fröhlich and Arnulf Kletzin (Darmstadt Technical University, DE). Hugo M. Botelho, Andreas Veith, Christian Bauer, Renate Fröhlich and Arnulf Kletzin purified the protein. Sequence analysis was performed by Arnulf Kletzin. CD spectra of oxidized and reduced Rieske and the T_m of the oxidized protein were determined by Sónia S. Leal and Vesna Prosinecki (ITQB). All other experiments were performed by Hugo M. Botelho.

3.1. Summary

Rieske proteins and Rieske ferredoxins are present in the three domains of life and are involved in a variety of cellular processes. Despite their functional diversity, these small iron-sulfur proteins contain a highly conserved all-beta fold, which harbors a [2Fe-2S] Rieske center. We have identified a novel subtype of Rieske ferredoxins present in hyperthermophilic Archaea, in which a two cysteine conserved SKTPCX$_{(2-3)}$C motif is found at the C-terminus. We establish that in the *Acidianus ambivalens* representative, RFd2, these cysteines form a novel disulfide bond within the Rieske fold, which can be selectively broken under mild reducing conditions insufficient to reduce the [2Fe-2S] cluster or affect the secondary structure of the protein, as shown by visible CD, absorption and ATR FT-IR. RFd2 presents all EPR, visible absorption and visible CD spectroscopic features of the [2Fe-2S] Rieske center. The cluster has a redox potential of +48 mV (25°C and pH 7) and a pK_a of 10.1 ± 0.2. These shift to +77 mV and 8.9 ± 0.3 respectively upon reduction of the disulfide. RFd2 has a melting temperature near the boiling point of water (T_m = 99°C, pH 7.0) but it becomes destabilized upon disulfide reduction (ΔT_m = -9°C, ΔC_m = -0.7 M GuHCl). This example illustrates how the incorporation of an additional structural element such as a disulfide bond in a highly conserved fold such as that of the Rieske domain, may fine-tune the protein for a particular function or for an increased stability.

3.2. Introduction

Iron sulfur (Fe-S) proteins are ubiquitous in all life domains and are involved in many fundamental cellular processes, from electron transfer, to catalysis, to regulation of gene expression. The simpler Fe-S proteins, despite their low structural complexity and small size (<150 amino acids), are able to harbor clusters with different nuclearities ([2Fe-2S], [3Fe-4S], [4Fe-4S]), or

even combinations of clusters, as in the di-cluster ferredoxins ([3Fe-4S][4Fe-4S] and [4Fe-4S][4Fe-4S]). In these proteins the inorganic moieties are generally found in a single domain, having rather conserved topologies [1]. The protein fold and the Fe-S clusters have an intertwined contribution on the structural stability of these proteins: whereas polypeptide scaffolding provides a relatively rigid and protective ligand environment for cluster binding, the chemical stability of the Fe-S cluster itself plays a key stabilizing effect over the structure.

The conserved folds of simple Fe-S proteins are in some cases decorated by additional structural elements that somehow broaden and shape the structural and functional landscape of these proteins. Examples include N- and C-terminal amino acid extensions within dicluster ferredoxins [2-3], insertion of a structural zinc site [4], and incorporation of a disulfide bond within the fold [5-7]. For instance, di-cluster [3Fe-4S][4Fe-4S] ferredoxins [3] contain a N-terminal extension, which modulates the protein stability either by hosting an additional His/Asp zinc center [4, 8] or by creating a stabilizing hydrophobic core [9]. Also, a disulfide bond has been shown to contribute significantly to the stability of the [2Fe-2S] plant-type ferredoxin from the hyperthermophilic bacterium *Aquifex aeolicus*, as its presence increases the melting temperature from 113 to 121°C [6]. One last example is given from the respiratory-type Rieske proteins, in which a strictly conserved disulfide in the immediate vicinity of the [2Fe-2S] cluster (within 5 Å) stabilizes the cluster binding loops and modulates the redox potential [10].

In this context, Rieske proteins are interesting examples of the functional and structural diversity within a family of proteins that share the same basic all-beta fold [1, 11]. These proteins have the distinct feature of harboring a [2Fe-2S] cluster coordinated by two cysteine and two histidine residues. Like in other Fe-S proteins, a considerable diversity is allowed at the level of the

primary structure without disturbing the structural topology whereas the type of amino acid side chains and the electrostatic environment found around the [2Fe-2S] Rieske cluster modulate the redox potential of the center. This results in a broad range of functional potentials, from around -150 mV for the soluble Rieske ferredoxins from the dioxygenase components [12], to up to +400 mV for the photosynthetic or respiratory Rieske domains [13-14]. Subsequently to a thorough phylogenetic analysis [11, 15], the diversity of Rieske proteins was considerably expanded, in particular in what concerns the soluble Rieske type ferredoxins, as many novel sequences were identified mostly within phyletically distant organisms from the Archaea domain. Subsequent analysis using complete genome data has allowed to establish the existence of evolutionarily related sequences within the *Thermoprotei*, an Archaeal class comprising hyperthermoacidophilic organisms belonging to the *Sulfolobus* and *Caldisphaera* genus [16]. The genes identified encode for proteins with the typical Rieske fingerprint, but with very little amino acid identity towards other Rieske proteins. Here we report a novel structural feature within the all-beta Rieske fold, which was identified upon studying one of these as yet uncharacterized archaeal Rieske ferredoxins from *Acidianus ambivalens* (EMBL FN557298). It consists of a disulfide bond within the Rieske fold, whose redox status influences the properties of the [2Fe-2S] center and simultaneously stabilizes the protein. This particular disulfide is a novel feature within Rieske proteins, and is unrelated to the disulfides described in respiratory and photosynthetic Rieske domains [17] and in the archaeal respiratory-type SoxF [5] and sulredoxin [18] proteins, which have distinct locations in the fold. This finding illustrates how minor structural modifications modulate protein function and stability and is key to rationalize the evolutionary adaption of simple Fe-S proteins to distinct conditions and to understand how relatively

simple folds may be fine tuned for a particular function or for an increased stability.

3.3. Materials and methods

3.3.1. Sequence comparison

Homologous Rieske amino acid sequences were extracted from the non redundant protein database using BLAST and *A. ambivalens* Rieske Ferredoxin 1 (Genbank 22204177) and Rieske ferredoxin 2 (RFd2), *Sulfolobus tokodaii* Sulredoxin and *S. solfataricus* ARF as seeds. All organisms were included in the search. We have considered the hits with E-values of less than 10^{-10} significant. Sequences were aligned using ClustalX [19], MAFFT (at http://align.bmr.kyushu-u.ac.jp/mafft/online/server/) [20] , GeneDoc [21], and CINEMA (http://www.bioinf.manchester.ac.uk/dbbrowser/CINEMA2.1/) [22]. Dendrograms were calculated using the MAFFT server and displayed using PhyloWidget (http://www.phylowidget.org/). The nucleotide sequence of the *rfd2* gene was submitted to the EMBL database and was assigned the accession number FN557298.

3.3.2. Chemicals

All reagents were of the highest grade available commercially. Guanidinium hydrochloride (GuHCl) was obtained from Promega and the accurate concentration of its solutions was determined by refractive index measurements. 8-anilino-1-naphthalenesulfonic acid (ANS), Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and 5,5′-Dithiobis(2-nitrobenzoic acid) (DTNB) were purchased from Sigma.

3.3.3. Cloning, expression and purification

Acidianus ambivalens DSMZ 3772 was grown aerobically and anaerobically according to published procedures [23]. The gene encoding the

Rieske ferredoxin 2 (RFd2) was PCR-amplified from *A. ambivalens* genomic DNA using the following N-terminal and C-terminal primers TATTT CATGA TTAAG ACAAT TCTTT ATGAA and AAACA ACTCG AGTGT ACCTT TTCTA GGTAA TTC, respectively. The N-terminal primer swapped the native presumable GTG start codon for an ATG. The 336 bp PCR product was digested with BspHI and XhoI and ligated into the NcoI/XhoI-digested vector pET28a (Novagen, now Merck -Biosciences, Darmstadt, Germany) thus adding two amino acid codons for the XhoI restriction site and a C-terminal His-Tag to the native protein (..LEHHHHHH). The protein was produced after IPTG induction at an OD^{600nm} of 0.6-1 in *E. coli* BL21 CodonPlus RIL DE3 cells (Stratagene, now Agilent Technologies) grown overnight at 37˚C in LB medium supplemented with 1/1000 vol. of a 100 mM ferric chloride in 100 mM citric acid solution. 40 g of *E. coli* cells were resuspended after harvesting in a 400 ml of 40 mM potassium phosphate (KPi) buffer pH 8 containing 300 mM NaCl and 10 mM imidazole. The cells were broken using a continous French press (Constant Systems Ltd., Daventry, UK) followed by two centrifugation steps (15 min at 15,000 x g and 60 min at 150,000 x g, respectively). 200 ml of the clear supernatant were applied to a 20 ml Ni-NTA column (Qiagen, Hilden, Germany) equilibrated with the same buffer and connected to an FPLC (Pharmacia, now GE Health Care, Freiburg, Germany). The column was washed with 3 volumes of 50 mM imidazole and finally eluted with 500 mM imidazole, both in the same buffer. The colored fractions were combined and dialyzed against 40 mM KPi pH 8. The sample was subsequently dialyzed against 70 mM Tris-HCl pH 7.7 150 mM NaCl and loaded to a calibrated Sephadex G-50 column (2 x 50 cm) previously equilibrated with the same buffer. Rieske eluted as a pure sample as verified by SDS-PAGE [24] and the typical UV-vis absorption spectrum. This pink stock solution was concentrated to around 1-2 mg/ml by ultrafiltration using a 5 kDa cut-off membrane (Amicon Ultra, Millipore), flash frozen at liquid

nitrogen temperature and stored at -20°C for subsequent studies. Protein was quantified using Bradford's method [25].

3.3.4. Disulfide reduction

The Rieske ferredoxin was purified in the oxidized state. Full reduction was achieved with excess ascorbate or sodium dithionite at room temperature. Spectra were obtained with 4.6 μM (0.06 mg/ml) (far UV CD) or 15.2 μM (0.2 mg/ml) (near UV/visible CD and absorption) protein concentration in 2 mM KPi pH 7. Low buffer concentrations are required to minimize absorption in the far UV region. Other reduced Rieske samples were obtained by incubating 7.6 μM protein with 5 mM DTT or 5 mM TCEP, during 1 h in 50 mM KPi pH 7 at room temperature.

3.3.5. Spectroscopic methods

UV/Vis spectra were recorded in a Shimadzu UVPC-1700 spectrometer equipped with cell stirring and a Julabo water bath-coupled thermostated cell support. Unless otherwise indicated, UV/Visible absorption spectra were recorded at room temperature. Circular dichroism (CD) measurements were performed on a Jasco J-815 spectrometer and fluorescence measurements were acquired on a Cary Varian Eclipse instrument, both possessing peltier-thermostated cell supports. Attenuated total reflectance Fourier transform-infrared (ATR FT-IR) spectra were acquired in a Bruker IFS 66/S spectrometer equipped with a nitrogen-cooled MCT detector using the water-thermostatized Harrick BioATRcell II cell. EPR spectra were acquired on a Bruker EMX spectrometer equipped with an ESR 900 continuous-flow helium cryostat at 10 K, 9.39 GHz microwave frequency, microwave power 2.4 mW and modulation amplitude 1 mT.

3.3.6. Conformational stability

The conformational stability of Rieske was assessed by performing temperature- or chemical-induced denaturation while following several

spectroscopic properties: i. visible absorption, which reports on the [2Fe-2S] cluster integrity; ii. tyrosine fluorescence, which reports on tertiary contacts; and iii. ANS fluorescence, which reports on the exposure of hydrophobic patches. Protein was prepared as 7.6 µM in 50 mM KPi pH 7. Data treatment was performed as described in [26].

For thermal-induced denaturation, we incubated the Rieske ferredoxin with no reductant, 5 mM TCEP or 5 mM DTT plus 0 or 2.5 M GuHCl at 25°C. Then, we performed 1°C/min temperature ramps (25-92°C) while measuring the UV/visible absorption spectrum or tyrosine fluorescence at 304 nm. Fluorimeter was set up to excite at 275 nm using 10 nm excitation and emission slits and 600 V PMT voltage. Refolding was assessed by equilibrating the sample at 4°C overnight (absorption) or 25°C for 5 minutes (fluorescence).

GuHCl-induced chemical denaturation curves were obtained by incubating the Rieske ferredoxin with 0 or 5 mM TCEP and different concentrations of GuHCl during at least 3h at room temperature. The denaturation curve corresponds to the normalized variation of fluorescence intensity at 304 nm or absorbance at 330 nm. The chemical denaturation was further dissected by incubating the same samples with 10 fold molar excess ANS for at least 30 min at room temperature. In this case, the folding status was probed by the ANS emission band (λ_{exc} = 350nm).

3.3.7. ATR FT-IR spectroscopy

For ATR FT-IR, the Rieske ferredoxin was concentrated by ultrafiltration to 0.8 mM (10.5 mg/ml) in 40 mM Tris-HCl pH 7 and then centrifuged 30 min at 12000 x g to remove aggregates. Spectra were acquired at 20°C and 4 cm^{-1} resolution with 0 or 5 mM TCEP after 3 h equilibration. The amide I band was deconvoluted by fitting its Fourier self deconvolution with Lorentzian curves centered at the second derivative maxima and minima. Band assignment was

performed using the reference data in [27]. The secondary structure was estimated by the Lorentzian integrals.

3.3.8. Redox potentiometry

The reduction potential of the [2Fe-2S] Rieske cluster was determined at 20°C and with 7.6 μM protein concentration in 50 mM KPi pH 7 with and without 5 mM TCEP. The redox state of the Rieske [2Fe-2S] cluster was probed by the absorbance change at 490 nm. The following redox mediators were used: 1,2-naphtoquinone-4-sulfonic acid (+215mV, 0.6μM), 1,2-naphtoquinone (+180mV, 0.6μM), trimethylhydroquinone (+115mV, 0.6 μM), phenazine methosulfate (+80mV, 0.6μM), 1,4-naphtoquinone (+60mV, 0.6 μM), 5-hydroxy-1,4-naphtoquinone (+30mV, 0.6μM), methylene blue (+11mV, 0.79μM), menadione (0mV, 0.6 μM), indigo tetrasulfonate (-30 mV, 0.4 μM), indigo trisulfonate (-70mV, 0.55 μM), 2,5-hydroxy-p-benzoquinone (-130mV, 0.23 μM), 2-hydroxy-1,4-naphtoquinone (-152mV, 0.4μM; all potentials are in respect to the standard hydrogen electrode). Additionally, catalase (0.002 mg/ml), glucose oxidase (1.6 U/ml) and glucose (0.86 mM) were added to ensure O_2-free conditions. The fully oxidized protein was titrated anaerobically with buffered sodium dithionite. The sample with 5 mM TCEP was incubated 3 h at room temperature, after which time the [2Fe-2S] center had been reduced. The oxidized center was regenerated by bubbling with air. UV-Visible absorption spectra were recorded during the spontaneous re-reduction of the cluster. Several oxidation-reduction cycles were performed and the reduction

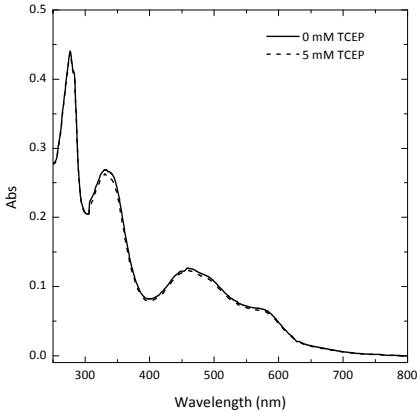


Figure 3.1 - UV-Visible absorption spectrum of RFd2 in the presence and absence of 5 mM TCEP. 15.2 μM (0.2 mg/ml) RFd2 was incubated in 50 mM KPi pH 7 in the presence or absence of 5 mM TCEP for at least 1.5 h at room temperature. The 5 mM TCEP sample spectrum remained unchanged for at least 5 h.

behavior did not change, despite the progressively lower re-reduction rates (~20-35 mV/min) due to reductant oxidation by oxygen. The titration curve includes datapoints from several reduction cycles. The midpoint reduction potential was determined by fitting a Nernst curve (*n*=0.7, 0 mM TCEP; *n*=1, 5 mM TCEP) to the normalized variation of absorption at 490 nm. The lower than unity value for the *n* value for a one electron acceptor center is within the experimental uncertainty and likely reflects a slight off-equilibrium in that set of measurements since the visible absorption spectrum of both samples is identical (Figure 3.1).

3.3.9. pH titration

The pK_a of the Rieske cluster was determined by investigating the pH-associated changes of the visible CD spectrum. 11.4 µM protein samples were prepared in a buffer solution composed of MES, HEPES, tricine, taurine and CAPS (10 mM each) and 50 mM NaCl with and without 5 mM TCEP poised at different pH values. The samples were equilibrated overnight at room temperature before measuring the near UV/visible CD spectrum at 25°C. One pK_a was determined by fitting the Hendersson-Hasselbalch formalism to the average normalized CD variation at 338, 378, 417, 434 and 485 nm (0 mM TCEP sample) or 338, 378, 462 and 485 nm (5 mM TCEP sample).

3.3.10. Iron quantification

Fe^{2+} was quantitated in Rieske ferredoxin samples using the chromogenic ferrous iron chelator TPTZ (2,4,6-tripyridyl-s-triazine) using a protocol adapted from [28]. Briefly, iron was released from Rieske by degradation in 0.9 M HCl. The protein fraction was removed by precipitation in 8% trichloroacetic acid. The remaining iron solution was buffered in 16% ammonium acetate and reduced with 0.9% hydroxylamine hydrochloride. The Fe^{2+} concentration was then determined by the rise in $Abs^{593\ nm}$ upon adding 320 µM TPTZ using a calibration with iron standards.

3.3.11. Thiol quantification

Four Rieske samples were prepared in 100 mM KPi pH 8: oxidized native, 5 mM TCEP reduced, denatured in 6 M GuHCl plus 2 mM EDTA, and denatured in 6 M GuHCl, 2 mM EDTA and 5 mM TCEP. Protein concentrations ranged from 23-69 μM and samples were incubated between 60 and 90 minutes. After incubation, TCEP-containing samples were desalted (HiTrap Desalting, GE Healthcare) in Argon-saturated 100 mM KPi pH 8-50 mM NaCl buffer. Then, the thiol content was determined by adding 2 mM DTNB and recording the quantitative formation of TNB^{2-} (ε^{412} = 14150 $M^{-1}.cm^{-1}$). All procedures were carried at room temperature.

3.4. Results and discussion

3.4.1. RFd2 is a representative of a subtype of Rieske ferredoxins

The protein used in this study, Rieske ferredoxin 2 (RFd2), was identified from a search on the *A. ambivalens* genomic data (A. Kletzin et al., unpublished) using Rieske protein sequences as seeds. Among the hits was also the previously characterized Rieske ferredoxin 1 [16]. The native RFd2 is a 106 amino acid protein with a theoretical mass of 12053 Da, and an isoelectric point of 8.9, whereas the length of the recombinant protein is 114 aa, including the His-Tag (13118 Da, pI = 8.6). Subsequent database analysis using BLAST has allowed the identification of 8 significant hits (E-value ≤ 10^{-17}) related to RFd2, all from unannotated hypothetical proteins from the *Sulfolobales* and *Caldisphaerales* orders within the Archaeal class of the *Thermoprotei*, which comprises hyperthermophilic organisms. Sequence alignment (Figure 3.2 and Supplementary Figure S3.1) and phylogenetic analysis (Figure 3.3) suggest that the latter cluster is a novel subgroup within the Rieske family. This group has two prominent features: one is a N-terminal truncation in a stretch which should include the β1 strand of the Rieske fold, and the other is the presence of an additional cysteine pair in a

$SKTPCX_{(2-3)}C$ motif at the C-terminal segment (Figure 3.2), in a position that maps into a loop region in the dioxygenase-type Rieske structures [12, 29]. The sequence conservation and close proximity of these additional cysteines is suggestive of a structural role, which may involve formation of a disulfide bond or involvement in coordinating an additional metal ion, such as zinc for example.

The *rfd2* gene was cloned, expressed in *E. coli* and the protein was purified to homogeneity using chromatographic methods. The resulting pink protein elutes as a monomer on gel filtration chromatography and migrates in 15%

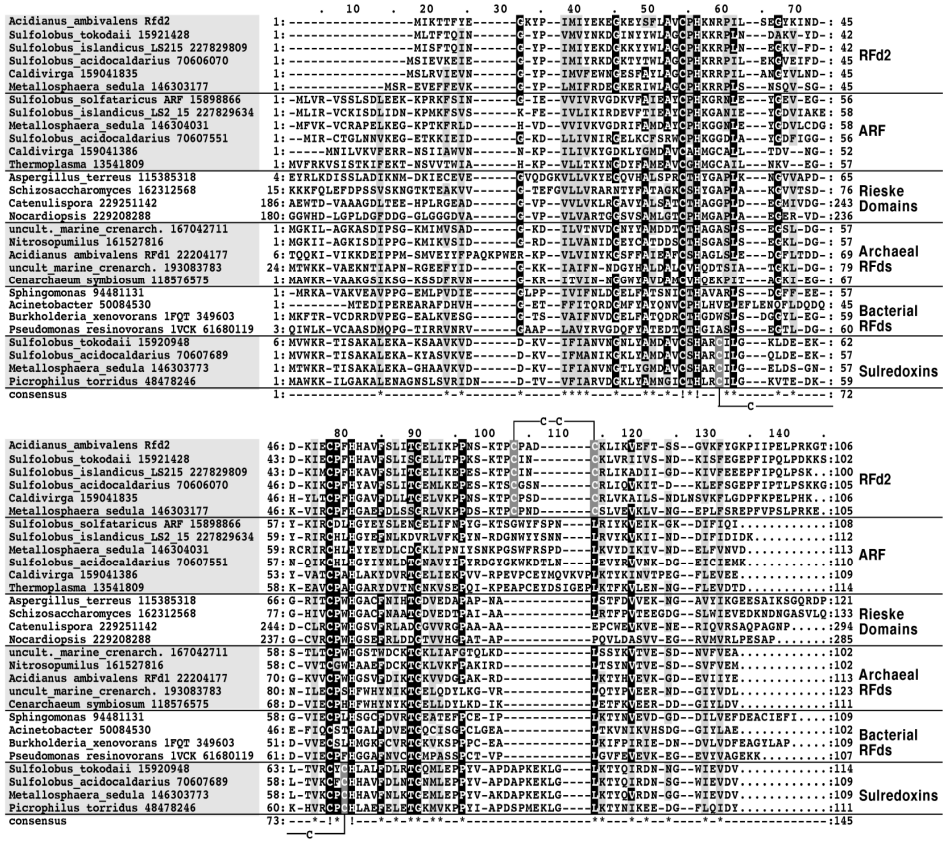


Figure 3.2 - Multiple sequence alignment of Rieske ferredoxins and ferredoxin domains. Rieske ferredoxins. RFd2, Rieske ferredoxin cluster RFd2; ARF, archaeal Rieske ferredoxins; lines with C-C denote the RFd2 disulfide bridge and the hypothetical disulfide bridge in the Sulredoxins. The latter is homologous to the disulfide from Rieske proteins from membrane-bound electron transport proteins (bc_1- and b_6f-complexes); numbers with species names denote the Genbank accession numbers.

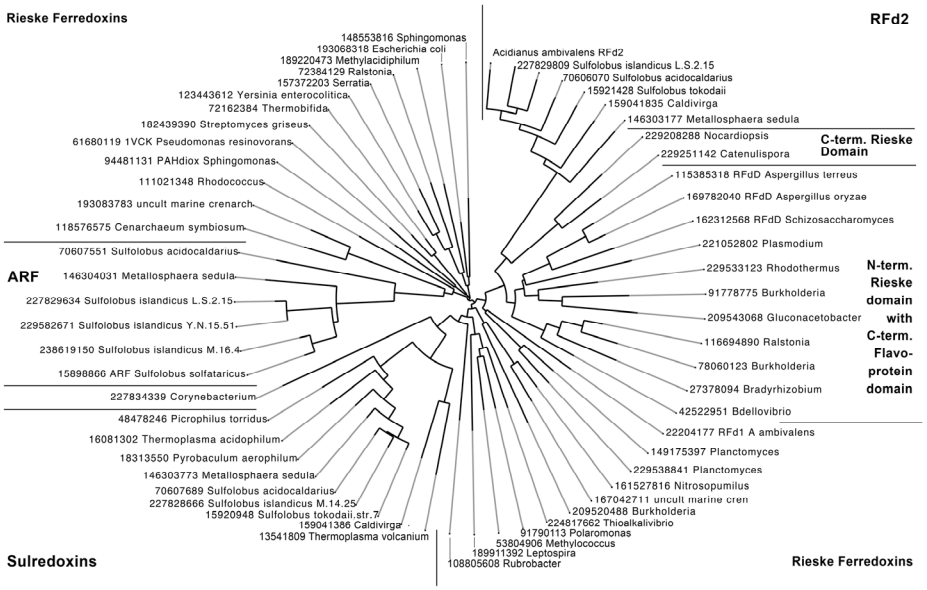


Figure 3.3 - Phylogenetic tree of Rieske ferredoxins calculated from the alignment shown in. Supplementary Figure S3.1. GI numbers from Genbank (http://www.ncbi.nih.gov); RFd2, cluster of RFds similar to the protein described in this work; ARF, archaeal Rieske ferredoxin.

SDS-PAGE gels as having 14.7 kDa (Figure 3.4). Iron quantitation using the TPTZ methods determined 2.4±0.1 (*n*=3) iron ions per molecule, which is compatible with the presence of an intact [2Fe-2S] cluster in the purified protein.

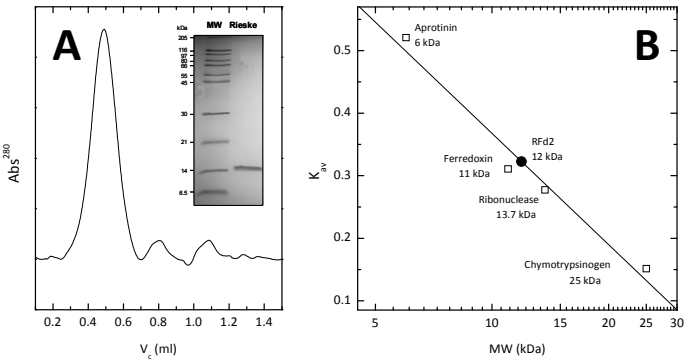


Figure 3.4 - Molecular weight and oligomeric state determination of purified RFd2. A: Gel filtration chromatogram (Sephadex G-50). Inset: 15% SDS-PAGE analysis of purified Rieske ferredoxin. The protein migrates as being 14.7 kDa. MW: molecular weight markers. B: molecular weight determination of the native protein using a calibration of the gel filtration column.

3.4.2. Spectroscopic characterization of the Fe-S cluster in RFd2

The presence of a [2Fe-2S] Rieske was clearly confirmed by subsequent biophysical analysis using UV-visible absorption, visible circular dichroism and EPR spectroscopies. The identity of the Rieske [2Fe-2S] center was unambiguously revealed by its characteristic EPR spectrum in the reduced form. It consists of a rhombic signal characterized by g values g_x = 1.80, g_y = 1.90 and g_z = 2.01 (Figure 3.5A). The average g-value is 1.905, typical for [2Fe-2S] clusters containing nitrogen ligands and significantly lower than the value around 1.96 for plant type [2Fe-2S] clusters. As expected for an integer spin system, no signal was observed in the as prepared, oxidized protein. The UV-visible absorption spectrum exhibits characteristic bands at 337, 456 and 570 nm, and a shoulder at 490 nm (Figure 3.5B). The circular dichroism spectrum in the visible region denotes positive peaks at 418 and 483 nm and negative peaks at 379, 436, 528 and 569 nm (Figure 3.5C). Altogether, the features observed by these two techniques are typical of Rieske clusters [5-6]. The Rieske center is readily reduced by dithionite (Figure 3.5BC, dotted lines) and partly reduced by ascorbate (Figure 3.5BC, dashed lines), indicating that the redox potential of the center has a positive value (see below).

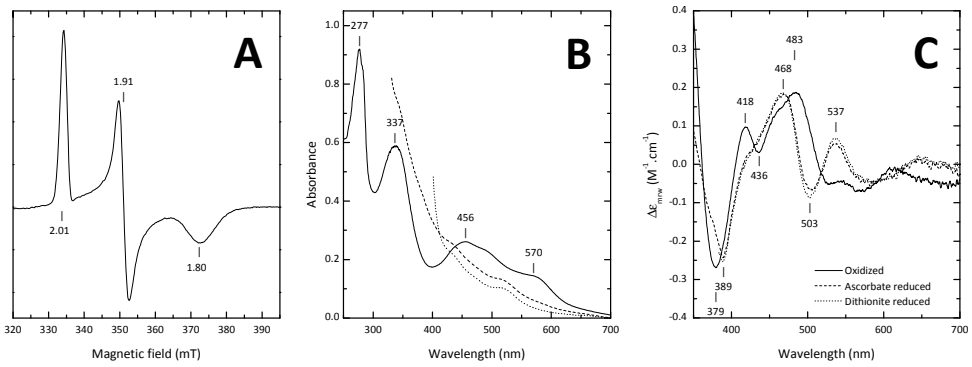


Figure 3.5 - Spectroscopic analysis of the Rieske cluster. Panel A: EPR spectrum of the dithionite reduced Rieske ferredoxin at 10 K, 9.39 GHz microwave frequency, microwave power 2.4 mW and modulation amplitude 1 mT. Panel B: UV/visible absorption spectrum of the oxidized (—) and ascorbate (----) and dithionite (···) reduced protein; Abs^{337nm}/Abs^{277nm} = 0.64. Panel C: visible CD spectrum of the oxidized and ascorbate and dithionite reduced protein.

3.4.3. The cysteines in the C-terminal SKTPCX$_{(2-3)}$C motif form a disulfide bridge

In order to address a possible role for the additional pair of cysteines (Cys-74 and Cys-78) from the SKTPCX$_{(2-3)}$C motif in disulfide formation, we have quantified the free thiols (-SH) in RFd2, under different conditions, using the DTNB assay (Figure 3.6).

The protein contains a total of four cysteine residues, two of which are involved in the coordination of the [2Fe-2S] Rieske cluster (Cys-29 and Cys-50). The as prepared protein was found to contain no free thiols (<0.1 SH.mol^{-1}), which shows that under these conditions cysteines are unavailable to the DTNB reagent (Figure 3.6, Scheme 1). On the other extreme, in a preparation in which the RFd2 was unfolded under reducing conditions and cluster disassembly was further promoted by the presence of EDTA, a total of 3.5±0.3 SH.mol^{-1} were quantified. This corresponds to all cysteines in the protein becoming available to the DTNB reagent (Figure 3.6, Scheme 4). On the other hand, under the same denaturing conditions, if the reductant is omitted, only 1.6±0.2 SH.mol^{-1} are quantified. This indicates that under these

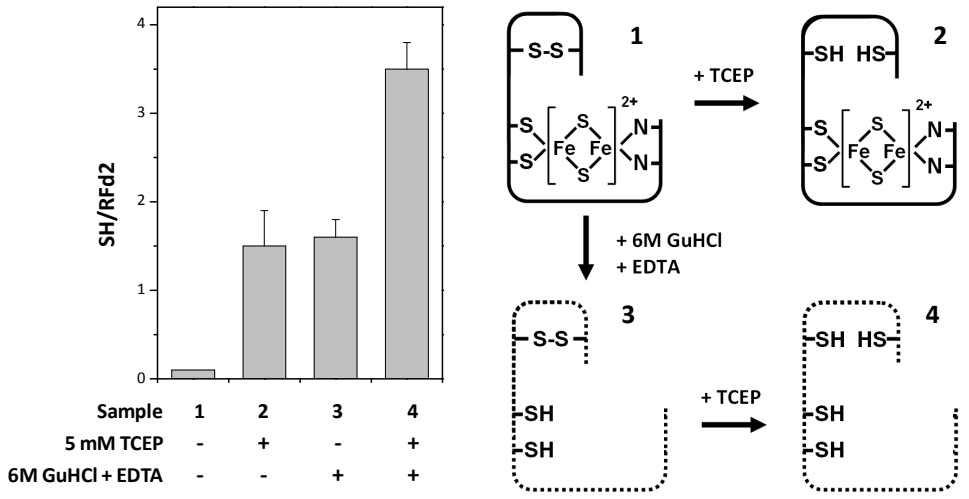


Figure 3.6 - Thiol quantitation and schematic representation of the different tested conditions.

conditions, the cluster binding cysteines become available whereas Cys-74 and Cys-78 are not, thus confirming that they are involved in a disulfide (Figure 3.6, Scheme 3). Additional control experiments were carried out in which the cluster binding cysteines and those from the SKTPCX$_{(2-3)}$C motif were selectively reduced. This was possible since we have observed that incubation of RFd2 with 5 mM TCEP is enough to reduce the disulfide bond (1.5±0.2 SH.mol^{-1}, Figure 3.5, Scheme 2) without reducing the Fe-S cluster, as the visible absorption spectrum remains unchanged even after 5 h of incubation (Figure 3. 1). This provides a tool to generate a protein form in which this disulfide is selectively broken (RFd2^{SH}). In summary, the cysteines within the C-terminal SKTPCX$_{(2-3)}$C motif are involved in a disulfide bridge, which is novel within the Rieske fold. The other types of disulfides known in Rieske proteins, further discussed below, are both found in respiratory Rieske domains, namely on the prototypic complex III domains [17] and on the SoxF protein from *Sulfolobus acidocaldarius* [5, 30].

3.4.4. The Rieske fold and secondary structure are not disrupted upon disulfide reduction

To fully characterize the role of the redox status of the disulfide bridge in the folding and conformation of RFd2, we have carried out a structural analysis using biophysical methods. Circular dichroism spectroscopy was inadequate for such a study as the high content in β sheets of the Rieske fold accounts for a relatively featureless spectrum in the far-UV region, with a minimum around 200 nm, typical of

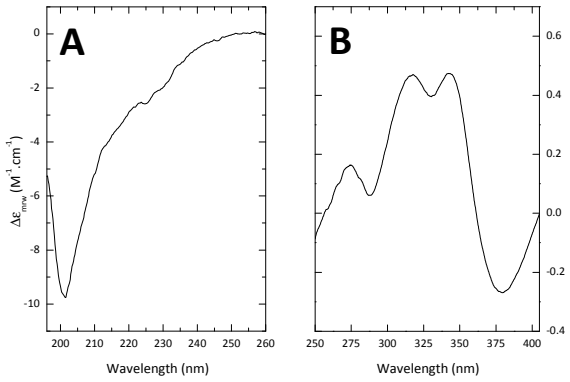


Figure 3.7 - Circular dichroism spectra of RFd2. Far-UV (A) and near-UV (B) regions.

proteins with high amounts of random coil and no significant amounts of α-helices, while showing the expected tertiary structure fingerprint on the near-UV region (Figure 3.7). However, the multiple vibrational modes associated with β structures produce highly informative infrared absorption spectra. For the purpose, we have used ATR FT-IR to identify and estimate the relative amounts of each type of secondary structure elements in RFd2 preparations with the disulfide in the oxidized ($RFd2^{SS}$) and in the reduced states ($RFd2^{SH}$). This was achieved by deconvoluting the amide I band and carrying out assignments according to established fingerprints. This analysis showed that the spectra are nearly identical in both preparations (Figure 3.8). The secondary structure of the oxidized protein (Figure 3.8A) is composed of 43% β-sheets, 25% random coil, 18% turns and 8% α-helices,

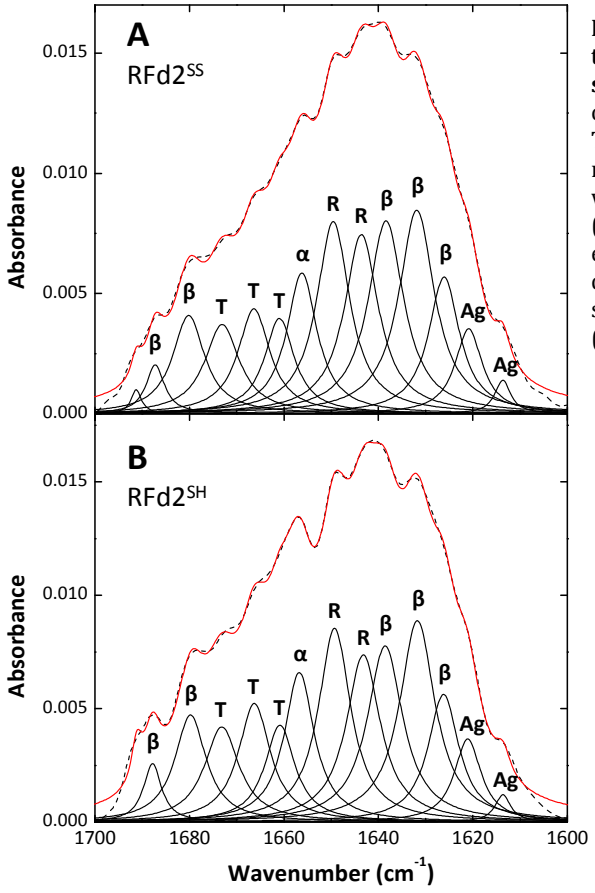


Figure 3.8 - ATR FT-IR analysis of the Rieske ferredoxin secondary structure. Self deconvoluted spectra of Rieske (black dashes) in 40 mM Tris pH 7 plus 0 mM TCEP (A) or 5 mM TCEP (B). The spectra were fitted with the sum of Lorentzian curves (red solid line). The assignment of each band (black solid lines) is depicted: β-sheets (β), unordered structures (R), α-helices (α), β-turns (T) and intermolecular β-sheets (Ag).

Table 3.1 - Assignment and relative contributions of the RFd2 amide I band secondary structure components, in the intact ($RFd2^{SS}$) and broken disulfide ($RFd2^{SH}$) forms.

Band $\tilde{\nu}$ (cm^{-1})	$RFd2^{SS}$	$RFd2^{SH}$	Assignment
	Area (%)		
1626	8	8	β-sheets
1632	14	14	
1638	13	12	
1680	6	6.5	
1687	2	2	
1691	0.5	1	
1644	12	11	Unordered structure (random)
1650	12	12	
1656	8	9	α-helices
1661	5.5	5	β-turns
1666	6.5	7	
1673	6.0	7	
1614	1	1	Intermolecular β-sheets (aggregation)
1621	4.5	4.5	

which is in a fair agreement with the calculations made in the available Rieske structures. Around 6% of the secondary structure can be accounted as intermolecular β-sheet contacts, which are usually associated with aggregation. These are, however, not significant in the overall structural content. The secondary structure changes brought about by disulfide reduction (Figure 3.8B) are marginal and are within the deconvolution error (Table 3.1). Therefore, reduction of the disulfide does not result in partial unfolding or in a loss of secondary structure.

3.4.5. The properties of the Rieske [2Fe-2S] cluster are influenced by the redox status of the disulfide

Redox potentiometry was used to determine the oxidation-reduction potential of the [2Fe-2S] Rieske cluster, using visible absorption to monitor the reduction state of the center during an anaerobic reductive titration using dithionite. The process was reversible and a redox potential of +48 mV (pH 7) was determined (Figure 3.9, open circles). Taking advantage of the fact that the disulfide bond can be broken while the Rieske cluster remains oxidized, we have tested if the redox properties of the cluster would be

influenced by the status of the disulfide. We have therefore carried out a redox titration on the disulfide-reduced RFd2 (RFd^{SH}) and we have observed that the apparent redox potential of the Rieske increases around 30 mV, to an E^0= +77 mV (pH 7) (Figure 3.7, closed circles). These redox potentials are intermediate between those of prototypic dioxygenase-type Rieske Ferredoxins ($E^0 \approx$ -180 mV, [31]) and respiratory and photosynthetic Rieske domains ($E^0 \approx$ +300 mV, [17]), and this wide range have been suggested to result from differences in solvent accessibility, hydrogen bonds and spatial distribution of ionizable side chains around the center, among different types of Rieske proteins [29, 32]. The status of the disulfide bond influences also the redox properties of the [2Fe-2S] cluster, as its reduction results in a +30 mV increase in the potential. This effect may arise from a change in the hydrogen bond network around the cluster as a result of disulfide reduction. Further investigation of these factors in RFd2 will be possible when a crystal structure becomes available, as the structural model of the protein produced by I-TASSER [33] does not allow such detailed analysis.

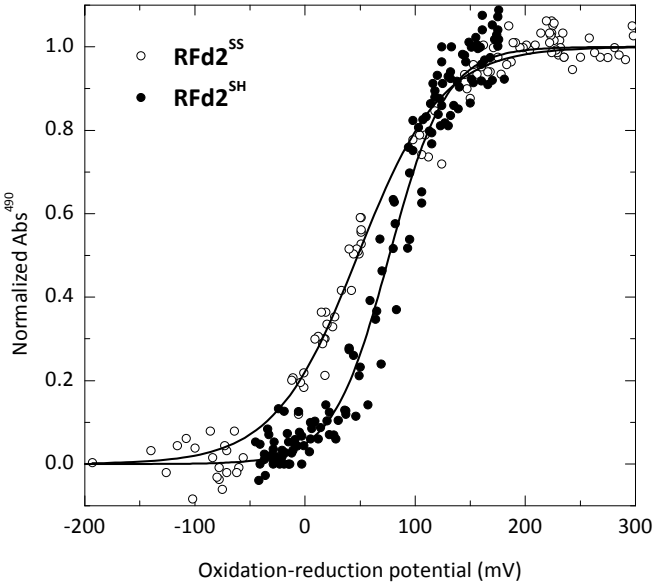


Figure 3.9 - Redox titration of the Rieske [2Fe-2S] center monitored by UV/visible absorption. The reduced fraction corresponds to the absorbance variation at 490 nm. The fits are Nernst curves for E^0 = +48 mV and n=0.7 (Oxidized Rieske) or E^0 = +77 mV and n = 1 (5 mM TCEP reduced).

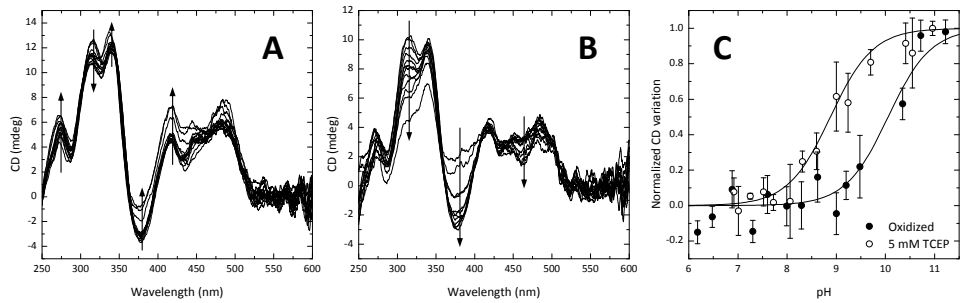


Figure 3.10 - Near UV/visible CD monitored pH titration of the Rieske ferredoxin. The pH titration was carried out on Rieske buffered in a complex buffer composed of MES, HEPES, tricine, taurine and CAPS (10 mM each) and 50 mM NaCl set at different pH values in the absence (A) or presence (B) of 5 mM TCEP. Arrows indicate the spectral variation at increasing pH values. pK_a values were extracted by fitting Handersson-Hasselbalch equations for a single protonation to the averaged normalized CD signal at 338, 378, 417, 434 and 485 nm (0 mM TCEP, ●) or 338, 378, 462 and 485 nm (5 mM TCEP, ○) (C).

We have also investigated the influence of the redox status of the disulfide on the pK_a values of the histidine ligands of the [2Fe-2S] cluster, which can be estimated from the analysis of the pH dependence of the visible CD spectrum of the Rieske center, according to a well established procedure [10]. We have used the variation of the visible CD spectra of RFd2, in the as prepared and in presence of 5 mM TCEP, from pH 6.2 to 11.2 (Figure 3.10).

The data could be fit with a single transition at a pK_a = 10.1 ± 0.2 (0 mM TCEP sample) or pK_a = 8.9 ± 0.3 (5 mM TCEP sample), which are intermediate between the values reported for *bc₁* Rieske proteins [34] and Rieske ferredoxins [32] in the oxidized $[2Fe\text{-}2S]^{2+}$ state. The determined values probably correspond either to a single or to an average of the two expected transitions from the two protonatable histidines, which could not be resolved. The decreased pK_a in the disulfide reduced state reflects the double protonation event following the disulfide cleavage and constitutes additional evidence to the proximity of the intervening cysteines to the [2Fe-2S] cluster which is responsible for the dependence of the cluster properties with the redox status of the disulfide. We have confirmed that the spectral changes observed are due to pH-dependent modifications at the [2Fe-2S] cluster and not by a perturbation in the secondary structure or folding, as the

far UV CD spectra in both assays were not altered at the pH extremes (not shown). The presence of the C-terminal hexahistidine could result in a shift of the pK_a values compared to the native, tag-less protein. However, this hypothetical perturbation is minimized because the N-terminus and the [2Fe-2S] cluster are located in opposite sides of the Rieske fold.

3.4.6. The stability of RFd2 is influenced by the redox state of the disulfide

A role of the disulfide in the protein conformational stability was investigated performing chemical and thermal unfolding experiments under conditions in which the disulfide was formed and broken. We have monitored the RFd2 folding using the intrinsic fluorescence of tyrosine residues as a probe of the conformational state of the protein, as the protein contains no tryptophan. The chemical stability was determined as a function of the concentration of guanidinium hydrochloride (GuHCl) in the disulfide oxidized and reduced states (Figure 3.11).

From the difference in the apparent midpoint denaturant concentrations (C_m=3.7 M for $RFd2^{SS}$ versus C_m=3.0 M for $RFd2^{SH}$) it is clear that the disulfide increases the chemical stability of the protein. We further explored an effect of this bond o n the thermal denaturation of the protein. As many other proteins from hyperthermophilic origin, RFd2 unfolds near or above the boiling point of water, as no transition is observed when the temperature is increased from 25 to 95°C. In order to overcome this problem, a series of thermal denaturation transitions were determined in the presence of GuHCl at different concentrations, all below that of the midpoint chemical denaturation (C_m). Under these conditions, the protein becomes destabilized by the presence of denaturant but the majority of the molecules are in the folded conformation thus allowing the determination of a apparent T_m at different GuHCl concentrations, with the disulfide intact and broken (Figure 11BC). Linearization plots can then be used to extrapolate the T_m in the

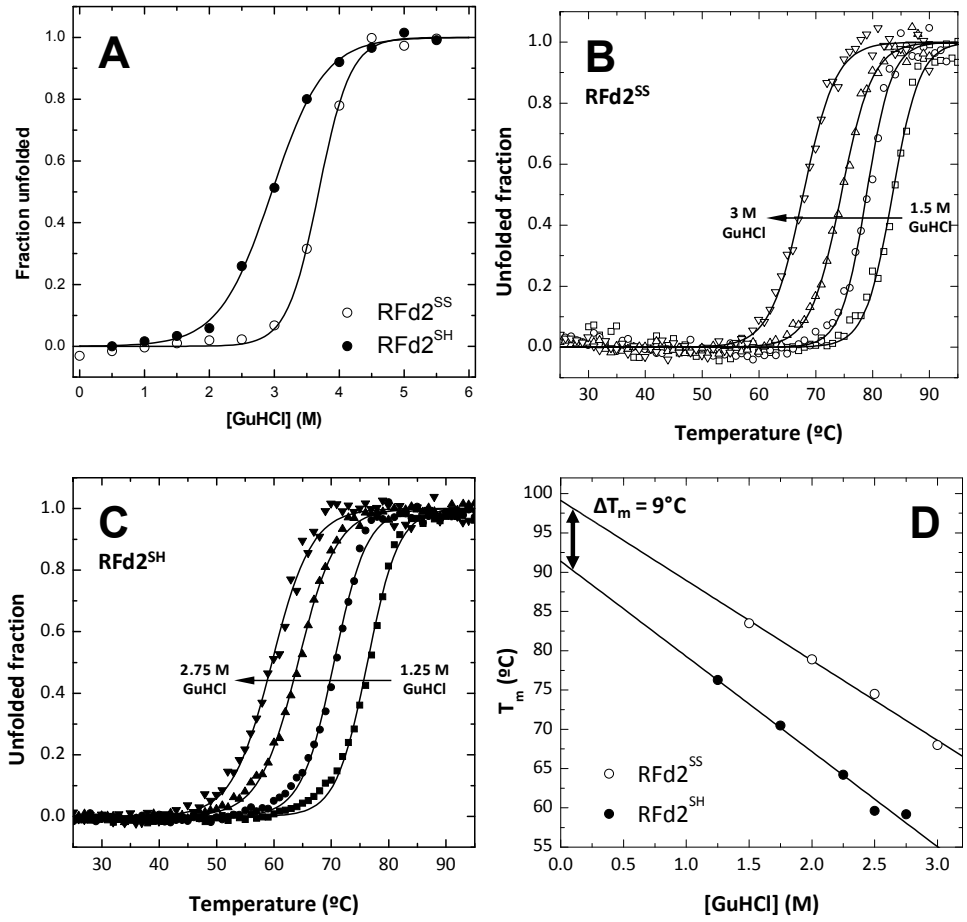


Figure 3.11 - Conformational stability of RFd2. Panel A: Chemical denaturation with GuHCl. Panel B and C: Determination of the midpoint denaturation temperature ( ) of the Rieske ferredoxin in the absence (B) and presence (C) of 5 mM TCEP, at different GuHCl concentrations. Panel D: Extrapolation plots for determination of the T_m in H_2O.

absence of denaturant, which is determined from the y-axis intercepts (Figure 3.11D). Using this strategy, we have determined that RFd2 has a $T_m^{H_2O}$ of 99°C, which is lowered to 91°C when the disulfide is broken. The integrity of the Rieske [2Fe-2S] cluster was also monitored upon thermal and chemical perturbation using visible absorption spectroscopy and identical results were obtained indicating that protein unfolding and cluster disintegration are intertwined events. Altogether, the chemical and thermal unfolding data

clearly show that the disulfide involving Cys-74 and Cys-78 does play an important contribution in structural stabilization.

3.5. Conclusions

Rieske proteins are present in the three domains of life, and are found to be involved in electron transfer processes associated with different cellular processes, including respiration, photosynthesis and dioxygenase systems [17, 35]. In recent years, studies on Rieske proteins from phylogenetically distant organisms, such as hyperthermophilic Archaea [15] have revealed proteins with particular features which enlarge the structural and functional diversity of Rieske domains. In fact, there are proportionally less Rieske proteins known in Archaea (~100) than in Eukarya (>450) or in Bacteria (>3500), as currently listed in Pfam [36]. Here we report the characterization of a new subtype of soluble Rieske ferredoxin of hyperthermophilic origin: in the essential this protein contains all typical fingerprints of Rieske domains in respect to the biophysical properties of its [2Fe-2S] cluster. However, it harbors a disulfide bond at the C-terminal region which has not yet been identified among the Rieske family, whose redox status affects the stability of the protein fold and of the iron-sulfur cluster, the redox properties of the latter and the pH-associated properties of the cluster ligands. Upon its disruption, a decrease in the chemical (ΔC_m = -0.7 M GuHCl) and thermal (ΔT_m = -9°C) stabilities is observed, an increase in the redox potential of the [2Fe-2S] center (ΔE^0 = +31 mV) and a decrease of the pK_a of the histidine ligands (1.2 units), likely as a result of a rearrangement of the electrostatic, hydrogen bonding and covalent network around the metal cluster environment.

Rieske proteins are believed to have the same structural fold, which consists of two all-beta subdomains: one conserved small domain which has a rubredoxin-like fold, and a larger domain which consists of 6 beta-stands packed in either sandwich of two 3-stranded sheets or closed barrel [17].

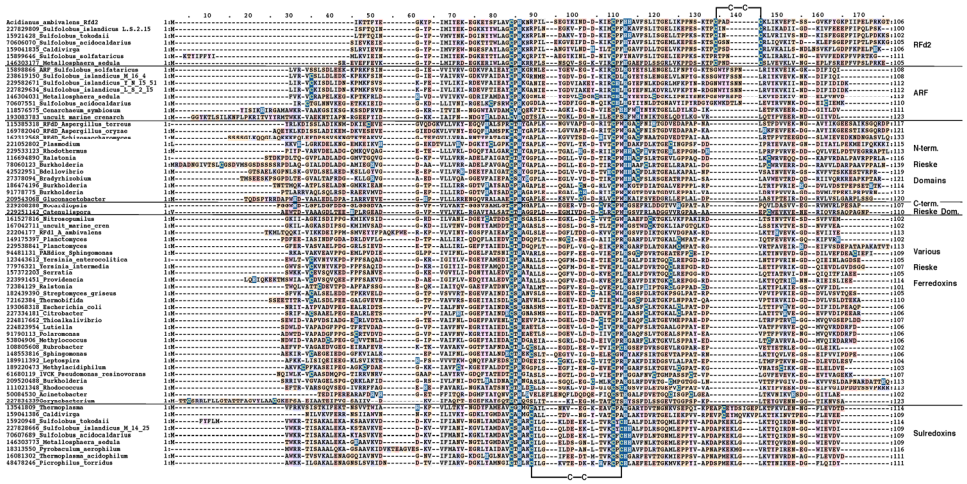
RFd2 typifies a new subtype of Rieske ferredoxins in which the Rieske fold is decorated by a novel disulfide bond which involves a cysteine pair located at the protein C-terminus, within a SKTPC$X_{(2-3)}$C motif. This motif appears as an insertion within a region that corresponds to a loop in dioxygenase Rieske proteins, and is present in five other sequences from Archaea, which share a rather high amino acid identity (~37-52%) among themselves, but not against other types of Rieske proteins (<15%). The disulfide bond now identified within this family is a novel structural feature in the Rieske fold, and is unrelated to the disulfides described in respiratory and photosynthetic Rieske domains [17] and in the archaeal respiratory-type SoxF [5] whose cysteines are intertwined within the FeS binding motif (CXHXXC…CPCH), and the one presumed to be present of sulredoxin [18].

The novel disulfide described within the conserved Rieske fold may underlie a stabilization strategy complementary to others found in highly thermostable proteins, which include extensive hydrogen bond and salt bridge networks, oligomerization, reduction of loop lengths, formation of tightly packed hydrophobic cores. In fact, the destabilization observed upon disulfide reduction is likely a consequence of the clipping effect exerted by the bond, as the motif containing the disulfide is inserted in a loop, which is also present in a shorter version in the dioxygenase Rieske ferredoxins. Insertion of disulfide bonds to cross-link different parts of the polypeptide chain with stabilizing effects is a frequent stabilizing strategy of local structures, and in fact thermophiles are predicted to contain a higher density of disulfide bonds than mesophiles [37]. Apart from the discussed examples within the Rieske family, other small iron-sulfur proteins also have a disulfide bond included on the protein structure. Those are the cases of the [3Fe-4S] ferredoxin II from *Desulfovibrio gigas* [7], the [4Fe-4S] ferredoxin from *Thermotoga maritima* and the plant-type [2Fe-2S] ferredoxin from

Aquifex aeolicus [6]. However, with the exception of the latter, the influence of disulfide on the protein stability has not been systematically investigated.

Overall, the work here reported illustrates how minor structural modifications modulate protein function and stability, and contributes to an understanding of how relatively simple folds may be fine-tuned. Further, the characterization of a subtype of Rieske proteins present in organisms rooting deeply on the tree of life is determinant to understand the evolutionary aspects of this fold, which is ubiquitous in nature.

3.6. Supplementary figure



Supplementary Figure S3.1- Multiple alignment of the Rieske ferredoxins and ferredoxin domains used for calculation of the dendrogram (Figure 3.3). RFd2, Rieske ferredoxin cluster RFd2; ARF, archaeal Rieske ferredoxins; lines with C-C denote the RFd2 disulfide bridge and the hypothetical disulfide bridge in the Sulredoxins. The latter is homologous to the disulfide from Rieske proteins from membrane-bound electron transport proteins (bc_1 and b_6f complexes); numbers with species names denote the Genbank GI numbers.

3.7. Acknowledgements

This work was supported by grants POCTI/QUI/3752, POCTI/QUI/45758 and POCI/BIO/58465 (to CMG) from the Fundação para a Ciência e Tecnologia (FCT/MCTES, Portugal). HMB (SFRH/BD/31126/2006), SSL (SFRH/BD/18653/2004 and SFRH/BPD/47477/2008/) and VP

(SFRH/BD/18746/2004) are recipients of fellowships from FCT/MCTES, Portugal.

3.8. References

1. Meyer J (2008) Iron-sulfur protein folds, iron-sulfur chemistry, and evolution. *J Biol Inorg Chem* **13**, 157-170
2. Gao-Sheridan HS, Pershad HR, Armstrong FA & Burgess BK (1998) Discovery of a novel ferredoxin from Azotobacter vinelandii containing two [4Fe-4S] clusters with widely differing and very negative reduction potentials. *J Biol Chem* **273**, 5514-5519
3. Gomes CM, Faria A, Carita JC, Mendes J, Regalla M, Chicau P, Huber H, Stetter KO & Teixeira M (1998) Di-cluster, seven-iron ferredoxins from hyperthermophilic Sulfolobales. *J Biol Inorg Chem* **3**, 449-507
4. Fujii T, Hata Y, Wakagi T, Tanaka N & Oshima T (1996) Novel zinc-binding centre in thermoacidophilic archaeal ferredoxins. *Nat Struct Biol* **3**, 834-837
5. Bönisch H, Schmidt CL, Schäfer G & Ladenstein R (2002) The structure of the soluble domain of an archaeal Rieske iron-sulfur protein at 1.1 A resolution. *J Mol Biol* **319**, 791-805
6. Meyer J, Clay MD, Johnson MK, Stubna A, Munck E, Higgins C & Wittung-Stafshede P (2002) A hyperthermophilic plant-type [2Fe-2S] ferredoxin from Aquifex aeolicus is stabilized by a disulfide bond. *Biochemistry* **41**, 3096-3108
7. Kissinger CR, Adman ET, Sieker LC, Jensen LH & LeGall J (1989) The crystal structure of the three-iron ferredoxin II from Desulfovibrio gigas. *FEBS Lett* **244**, 447-450
8. Kojoh K, Matsuzawa H & Wakagi T (1999) Zinc and an N-terminal extra stretch of the ferredoxin from a thermoacidophilic archaeon stabilize the molecule at high temperature. *Eur J Biochem* **264**, 85-91
9. Rocha R, Leal SS, Teixeira VH, Regalla M, Huber H, Baptista AM, Soares CM & Gomes CM (2006) Natural domain design: enhanced thermal stability of a zinc-lacking ferredoxin isoform shows that a hydrophobic core efficiently replaces the structural metal site. *Biochemistry* **45**, 10376-10384
10. Merbitz-Zahradnik T, Zwicker K, Nett JH, Link TA & Trumpower BL (2003) Elimination of the disulfide bridge in the Rieske iron-sulfur protein allows assembly of the [2Fe-2S] cluster into the Rieske protein but damages the ubiquinol oxidation site in the cytochrome bc1 complex. *Biochemistry* **42**, 13637-13645
11. Schmidt CL & Shaw L (2001) A comprehensive phylogenetic analysis of Rieske and Rieske-type iron-sulfur proteins. *J Bioenerg Biomembr* **33**, 9-26
12. Colbert CL, Couture MM, Eltis LD & Bolin JT (2000) A cluster exposed: structure of the Rieske ferredoxin from biphenyl dioxygenase and the redox properties of Rieske Fe-S proteins. *Structure* **8**, 1267-1278
13. Denke E, Merbitz-Zahradnik T, Hatzfeld OM, Snyder CH, Link TA & Trumpower BL (1998) Alteration of the midpoint potential and catalytic activity of the rieske iron-sulfur protein by changes of amino acids forming hydrogen bonds to the iron-sulfur cluster. *J Biol Chem* **273**, 9085-9093
14. Holton B, Wu X, Tsapin AI, Kramer DM, Malkin R & Kallas T (1996) Reconstitution of the 2Fe-2S center and g = 1.89 electron paramagnetic resonance signal into overproduced Nostoc sp. PCC 7906 Rieske protein. *Biochemistry* **35**, 15485-15493
15. Schmidt CL (2004) Rieske iron-sulfur proteins from extremophilic organisms. *J Bioenerg Biomembr* **36**, 107-113
16. Kletzin A, Ferreira AS, Hechler T, Bandeiras TM, Teixeira M & Gomes CM (2005) A Rieske ferredoxin typifying a subtype within Rieske proteins: spectroscopic, biochemical and stability studies. *FEBS Lett* **579**, 1020-1026
17. Link TA (2001) Fe-S Rieske center. In *Handbook of Metalloproteins* (Bode W, Cygler V & Messerschmidt A, eds). John Wiley and Sons, Chichester.
18. Iwasaki T, Imai T, Urushiyama A & Oshima T (1996) Redox-linked ionization of sulredoxin, an archaeal Rieske-type [2Fe-2S] protein from Sulfolobus sp. strain 7. *J Biol Chem* **271**, 27659-27663
19. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948
20. Katoh K & Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* **9**, 286-298

21. Nicholas KB, Nicholas HBJ & Deerfield DWI (1997) GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEWNEWS* **4**, 14
22. Parry-Smith DJ, Payne AW, Michie AD & Attwood TK (1998) CINEMA--a novel colour INteractive editor for multiple alignments. *Gene* **221**, GC57-63
23. Zillig W, Yeats S, Holz I, Bock A, Gropp F, Rettenberger M & Lutz S (1985) Plasmid-related anaerobic autotrophy of the novel archaebacterium Sulfolobus ambivalens. *Nature* **313**, 789-791
24. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
25. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254
26. Pace CN, Shirley BA & Thomson JE (1990) Measuring the conformational stability of a protein. In *Protein structure - a practical approach* (Creighton T, ed, pp. 311-330. IRL Press, Oxford.
27. Barth A & Zscherp C (2002) What vibrations tell about protein. *Quarterly Reviews of Biophysics* **35**, 369-430
28. Fischer DS & Price DC (1964) A Simple Serum Iron Method Using the New Sensitive Chromogen Tripyridyl-S-Triazine. *Clin Chem* **10**, 21-31
29. Brown EN, Friemann R, Karlsson A, Parales JV, Couture MM, Eltis LD & Ramaswamy S (2008) Determining Rieske cluster reduction potentials. *J Biol Inorg Chem* **13**, 1301-1313
30. Schmidt CL, Anemuller S & Schafer G (1996) Two different respiratory Rieske proteins are expressed in the extreme thermoacidophilic crenarchaeon Sulfolobus acidocaldarius: cloning and sequencing of their genes. *FEBS Lett* **388**, 43-46
31. Kimura S, Kikuchi A, Senda T, Shiro Y & Fukuda M (2005) Tolerance of the Rieske-type [2Fe-2S] cluster in recombinant ferredoxin BphA3 from Pseudomonas sp. KKS102 to histidine ligand mutations. *The Biochemical journal* **388**, 869-878
32. Zu Y, Couture MM, Kolling DR, Crofts AR, Eltis LD, Fee JA & Hirst J (2003) Reduction potentials of Rieske clusters: importance of the coupling between oxidation state and histidine protonation state. *Biochemistry* **42**, 12400-12408
33. Wu S, Skolnick J & Zhang Y (2007) Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol* **5**, 17
34. Link TA (1994) Two pK values of the oxidized 'Rieske' [2Fe-2S] cluster observed by CD spectroscopy. *Biochimica et Biophysica Acta* **1185**, 81-84
35. Bugg TD & Ramaswamy S (2008) Non-heme iron-dependent dioxygenases: unravelling catalytic mechanisms for complex enzymatic oxidations. *Current opinion in chemical biology* **12**, 134-140
36. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, et al. (2008) The Pfam protein families database. *Nucleic Acids Res* **36**, D281-288
37. Beeby M, O'Connor BD, Ryttersgaard C, Boutz DR, Perry LJ & Yeates TO (2005) The genomics of disulfide bonding and protein stabilization in thermophiles. *PLoS Biol* **3**, e309

4. Dissecting the mechanism of molten globule formation by *Acidianus ambivalens* ferredoxin: an FT-IR study

Ferredoxin was purified by Sónia S. Leal (ITQB).

4.1. Summary

Acidianus ambivalens ferredoxin is a model thoroughly studied in respect to the biophysical characterization of the interplay of its iron-sulfur (FeS) clusters and zinc ions with protein conformation and stability. The molten globule state where ferredoxin rests after thermal denaturation at pH 2.5 constitutes a proposed candidate for the template conformation where FeS clusters are incorporated during protein maturation. Here we report a structural analysis of the ferredoxin molten globule. By performing a temperature-dependent FT-IR study of the denaturation of ferredoxin at pD 2.5 and 12 we have described protein denaturation events and molten globule formation in terms of the variation of secondary structure. We have identified a common structural component spectrum describing ferredoxin's core structure and thermal denaturation at acidic and basic pD. This study also showed that cold unfolding can be accessible to the apo states of ferredoxin, opening new possibilities for the use of this protein as a model for the study of this phenomenon.

4.2. Introduction

Iron-sulfur (FeS) clusters are inorganic protein cofactors occurring in the three domains of life which are required for biological activity, functional regulation and conformational stabilization. The [3Fe-4S][4Fe-4S] ferredoxin (Fd) from *Acidianus ambivalens* (Figure 2.10) constitutes a system with unique properties for the analysis of the interplay between FeS clusters, protein stability and conformational states. First, its primordial $(\beta\alpha\beta)_2$ fold [1] and cofactors [2] along with its high thermostability (T_m = 122°C at pH 7.0, [3]) should provide insights into basic protein stabilization strategies. Second, the core fold is expanded by a 30 amino acid N-terminal extension containing a structural Zn^{2+} site which can act as an additional stabilization factor [4]. Third, a high resolution X-ray crystal structure which includes the

intact FeS clusters is available (Figure 2.11) [5]. Finally, multiple intrinsic spectroscopic probes have allowed a comprehensive description of the folding and stability properties of this protein [3, 6-9].

Recently it has been shown that the *A. ambivalens* ferredoxin forms a stable molten globule state at pH 2.5 *in vitro* [7]. Acidification yields a protein conformation in the holo state with a native-like conformation, which upon thermal denaturation sits in an apo molten globule state with compactness identical to that of the native conformation, similar secondary structure content, no tertiary contacts and enhanced exposed hydrophobic surfaces. The stability of the molten globule was shown by its ability to unfold reversibly upon thermal or chemical denaturation. The accessibility of a stable, partially structured conformation is especially relevant in the context of FeS proteins. FeS clusters are required for correct folding as they are placed in the hydrophobic core of the $(\beta\alpha\beta)_2$ fold [5]. Although excess iron (Fe^{3+}) and sulfide (S^{2-}) can spontaneously be incorporated in ferredoxins *in vitro* resulting in cluster assembly [10-11], the millimolar concentrations required are not physiological and thermally unfolded ferredoxin does not reincorporate the disintegrated clusters [6]. FeS clusters biosynthesis is a tightly regulated process wherein *de novo* assembled clusters are thought to be transferred from scaffold proteins to the corresponding apo proteins, a process likely involving some binding promiscuity due to the large number of FeS acceptors. In fact, the IscU scaffold protein is a molten globule-like protein [12] and such states have been surmised to be the likely conformation for the acceptor proteins [7].

The high β-sheet content of *A. ambivalens* ferredoxin [5] is an intrinsic probe that can be used to obtain structural data using Fourier transform infrared spectroscopy (FT-IR), a technique highly sensitive to this type of secondary structures [13]. We have built on the knowledge on the stability and formation of the molten globule of ferredoxin by performing FT-IR

monitored temperature denaturation assays at pD 2.5. Equivalent thermal denaturation experiments at pD 12 – where ferredoxin unfolds irreversibly to a non molten globule conformation – were carried out as a control. By deconvoluting temperature-resolved spectral data, we have quantified secondary structure changes taking place during holo ferredoxin unfolding, molten globule formation and during reversible molten globule unfolding. Overall, our results allow proposing a model describing the main conformation changes occurring in a temperature-dependent manner.

4.3. Materials and methods

4.3.1. Ferredoxin and chemicals

Acidianus ambivalens ferredoxin (Fd) was purified as described before [14] and concentrated to 1.09 mM (12.5 mg/ml) by ultrafiltration. Then, the solvent was evaporated under a flow of dry nitrogen and replaced by 20 mM glycine or 50 mM potassium phosphate (KPi) in D_2O. pD was adjusted to 2.5 or 12, respectively, using the formula pD = pH + 0.4 (at 25°C) [15]. Any undisolved solutes – which were not visually apparent – were pelleted by centrifuging 25 minutes at 12000 g at 4°C. Final protein concentration was 1.1 mM (12 mg/ml). All chemicals were of the highest grade commercially available.

4.3.2. Thermal denaturation

The thermal denaturation of ferredoxin was monitored by Attenuated Total Reflectance Fourier Transform Infrared spectroscopy (ATR FT-IR) using a Bruker IFS 66/S spectrometer equipped with a nitrogen-cooled MCT detector using the thermostatized Harrick BioATR cell II and recording spectra in the amide I (1600-1700 cm^{-1}) and amide II (1500-1600 cm^{-1}) regions. Temperature was modified in the 20-94°C range in a discontinuous fashion, according to Figure 4.1: spectra were accumulated during 1 minute (97 accumulations), temperature was raised 2°C during around 45 seconds

and sample was equilibrated 15 seconds. Overall temperature change rate was 1.1°C/min. Spectral resolution was 2 cm^{-1}, scanner velocity was 20.0 kHz and aperture was 12 mm. To study the molten globule state forming at pD 2.5 after thermally denaturing holo ferredoxin, the sample was re-equilibrated at 20°C and a second temperature ramp was performed immediately with the same sample. The same strategy was employed with the pD 12 sample for comparative purposes. To determine midpoint denaturation temperatures (T_m), the second derivative of FT-IR spectra were calculated and sigmoidal curves were fitted to the temperature trend of local maxima and minima. Alternatively, the normalized variation at specific wavenumbers was fitted to sigmoidal curves or the sum of two sigmoidal curves. Pre- and post-transition baselines were corrected assuming linear dependence on temperature [16]. Assignment of secondary structure elements to wavenumbers was carried out using published reference values [13].

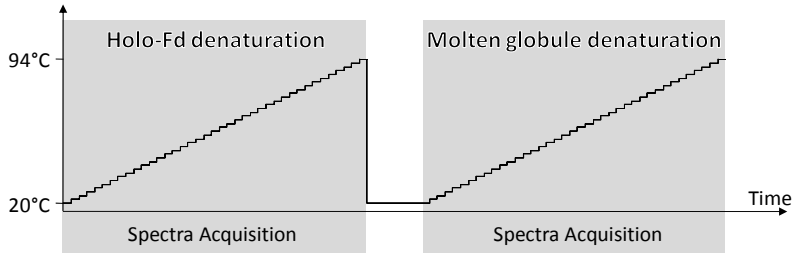


Figure 4.1 – Outline of the temperature denaturation experiments on *A. ambivalens* ferredoxin. Protein solutions at pD 2.5 or pD 12 were sequentially submitted to two 20-94°C temperature ramps (rate = 1.1°C/min) while ATR FT-IR spectra were recorded in a temperature-dependent manner.

4.3.3. Singular value decomposition

The series of spectra obtained during each temperature ramp were deconvoluted using the singular value decomposition (SVD) algorithm implemented in Matlab [17]. This allowed identifying spectral components changing their intensity as temperature was varied as well as their temperature variation profiles.

4.4. Results

We have studied the thermal denaturation of *Acidianus ambivalens* ferredoxin (Fd) by recording ATR FT-IR spectra in the amide I (1600-1700 cm^{-1}) and amide II (1500-1600 cm^{-1}) bands during thermal denaturation. These spectral regions are sensitive to protein secondary structure. Most importantly, the amide I band arises due to quantitative contributions from distinct secondary structure features [13]. This implies that absorption variation at specific wavenumbers is associated with secondary structure conversion. To increase sensitivity we have used deuterated buffer solutions. Deuteration shifts the water absorption peak at the amide I/II region and increases the relative contribution of protein absorption to the recorded spectra. Fd was prepared at pD 2.5 or 12 and each preparation was subjected to two sequential temperature ramps (T-ramps, Figure 4.1). Our goal is to analyze the molten globule formation occurring after thermal denaturation at pH 2.5 and its conformational and stability properties. pD 12 preparations are controls, as under these conditions ferredoxin denatures to a non molten globule state. The spectra recorded during temperature ramps are shown in Figure 4.2. This extreme of pH had to be used as a control rather than pH 7 as Fd is a hyperstable protein whose midpoint unfolding temperature (T_m) is above 100°C in the range pH 4.5 to pH 9.5 [9]. At pH 7, ferredoxin thermal unfolding below the boiling point of water requires the use of chemical denaturants such as guanidinium hydrochloride [9] or guanidinium thiocyanate [6], whose contribution to infrared absorption is significant, because of the high concentrations required.

During thermal denaturation at pD 2.5, a band at 1625 cm^{-1} develops (Figure 4.2A). Absorption at this wavenumber is typical for β-sheet structures and is frequently an indication of aggregation. Calorimetric data had previously indicated this possibility [7]. During the second T-ramp (Figure 4.2C), which corresponds to the denaturation of the molten globule,

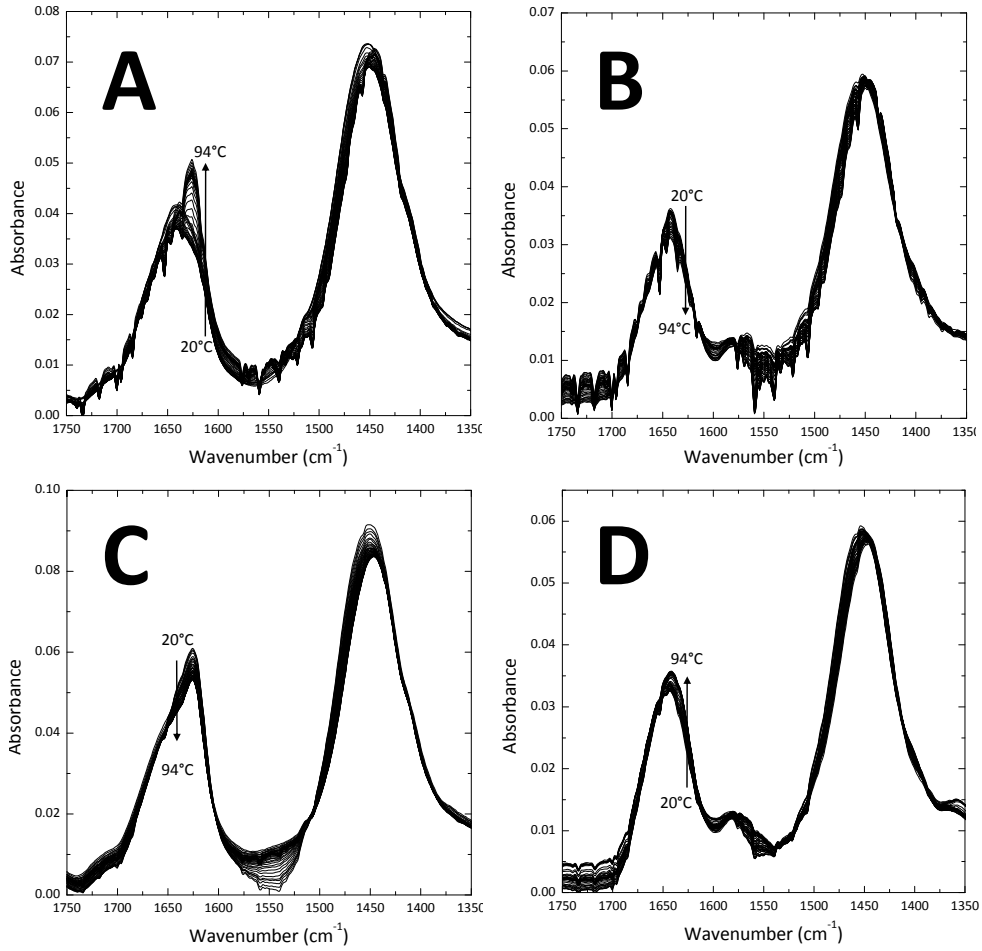


Figure 4.2 – ATR FT-IR monitored temperature ramps of *Acidianus ambivalens* ferredoxin. Spectra obtained along the 20-94°C temperature range at pD 2.5 (A, C) or pD 12 (B, D). The ferredoxin samples at either pD were subjected to two successive temperature ramps. (A, B) first temperature ramp. (C, D) second temperature ramp.

no significant spectral changes occur. This indicates that no significant secondary structure changes take place and is in agreement with far UV CD data [7]. Thermal denaturation at pD 12 does not yield significant spectral changes (Figure 4.2CD). The noise in the spectra acquired at 94°C on the first T-ramp precludes more detailed analyses. The major spectral variations can be better assessed in Figure 4.3. In all cases, absorption variations in the amide II region are not discussed in structural terms because of the significant contribution of side chains in this region.

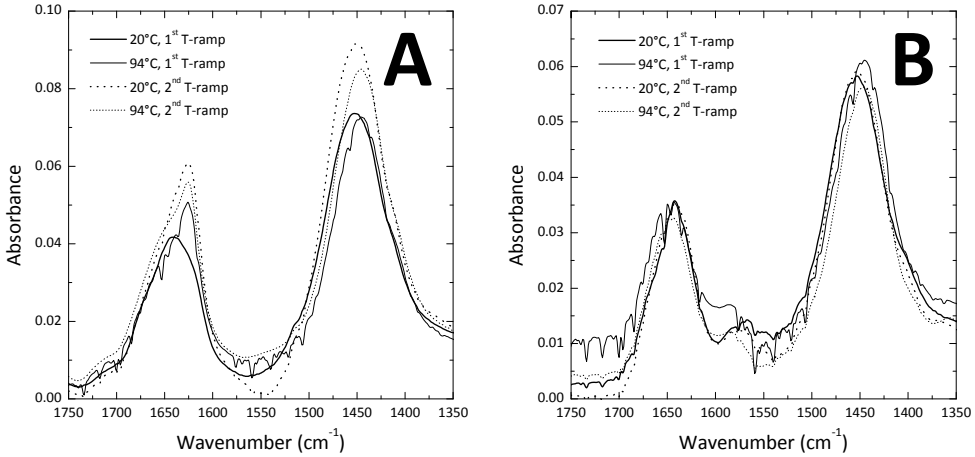


Figure 4.3 – Amide I and amide II bands before and after thermal denaturations. (A) pD 2.5. (B) pD 12.

The temperature-induced increase in β-sheet structures (1625 cm^{-1}, Figure 4.2A) can be used to quantify protein denaturation. We have used the absorption variation at 1625 cm^{-1} (pD 2.5) or 1643 cm^{-1} (amide I maximum absorption wavenumber at pD 12) to determine the overall midpoint denaturation temperature (T_m, Figure 4.4).

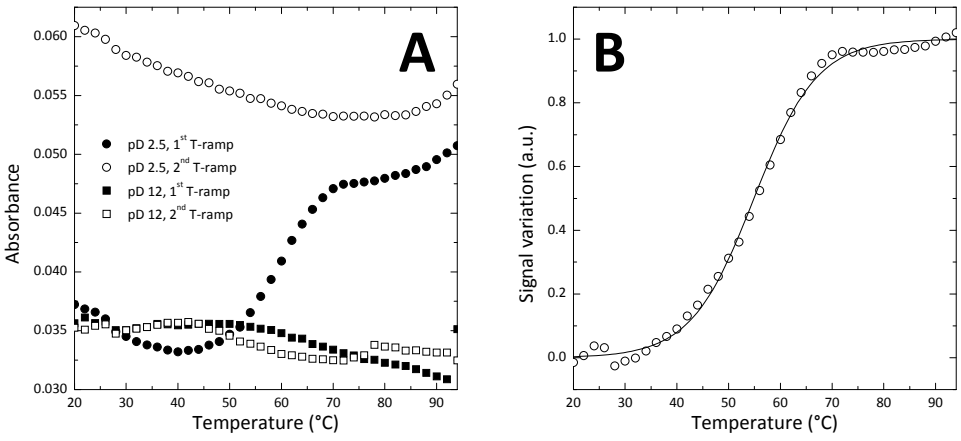


Figure 4.4 – Thermal denaturation profiles extracted from absorbance changes in the amide I band. (A) Temperature-dependent absorbance changes at 1625 cm^{-1} (pD 2.5) or 1643 cm^{-1} (pD 12). (B) Determination of the T_m value from the absorbance change at 1625 cm^{-1} for the first temperature ramp of ferredoxin at pD 2.5. Raw data (A) was compensated for the slopy pre- and post-transition baselines (see materials and methods).

Only the first T-ramp at pD 2.5 yielded a cooperative transition at T_m = 55°C. Due to the intrinsic complex nature of the amide I band, resulting from multiple contributions in this region, difference spectroscopy analysis provided limited information. Alternatively, we have analyzed the second derivatives of the temperature-dependent FT-IR spectra. This technique enhances the fine structure of spectra and enhances low intensity overlapping bands which appear as second derivative minima [18]. By performing this analysis to our set of thermal denaturation experiments, we were able to accurately describe the absorption variation at specific wavenumbers (Figure 4.5). Further, by assigning each wavenumber to a particular type of secondary structure, we were able to interpret in detail the structural changes taking place during thermal denaturation (Table 4.1).

Although all temperature profiles indicated in Table 4.1 have structural correlates, direct connections can be established with the profiles derived from negative second derivative peaks (bold in Table 4.1), which correspond to positive peaks in the original spectra. The T_m values for the first thermal denaturation at pD 2.5 group around two distinct values: 43-47°C (for α-helices, coil and turns) and 53-59°C (for β-sheets). Denaturation of β-sheets occurs at the same temperature where aggregation and iron-sulfur cluster loss is reported to occur [7], whereas the denaturation events at lower temperatures were not detected by other spectroscopic techniques. Concentration-dependent effects can be excluded *a priori* as previous biophysical studies were reported in a broad range of concentrations, from the micromolar (*e. g.* circular dichroism, [7]) to the millimolar (*e. g.* NMR, [8]) ranges. The second T-ramp, which corresponds to the denaturation of the molten globule formed after the first thermal denaturation, yields complex denaturation curves arising from the lower signal to noise ratio, as most secondary structure is lost upon molten globule formation (low secondary

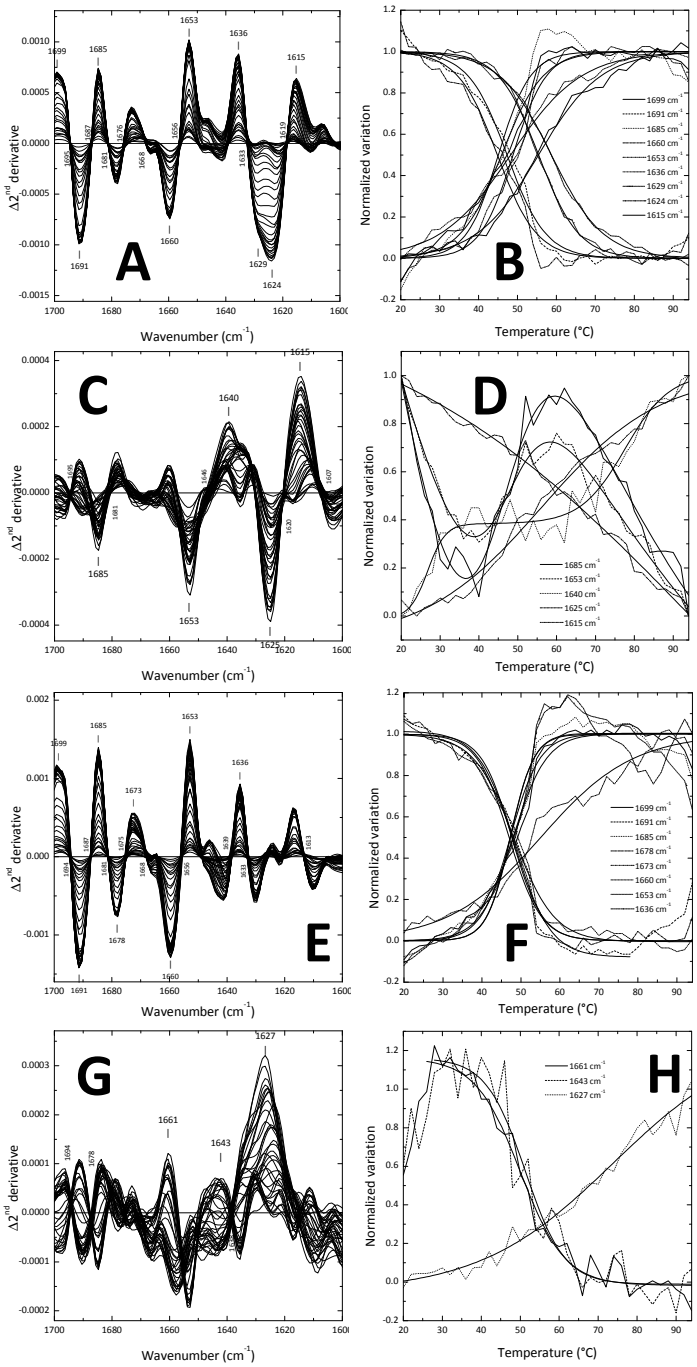


Figure 4.5 – Second derivative FT-IR spectra for the thermal denaturation of ferredoxin
(Left column) Second derivative spectra. (Right column) Normalized temperature-dependent second derivative changes at local maxima and minima. T_m values were extracted from fitting single or the sum of two sigmoidal curves to raw data (Table 4.1). (A-B) pD 2.5, 1st T-ramp; (C-D) pD 2.5, 2nd T-ramp; (E-F) pD 12, 1st T-ramp; (G-H) pH 12, 2nd T-ramp.

structure content is a structural feature of molten globule states). At pD 12 denaturation is highly cooperative and involves all secondary structure elements around 50°C. Denaturation is mostly irreversible as judged by the lower absolute values of second derivatives. In the second T-ramp all contributions are not arising from spectral peaks (*i. e.* are positive second derivative peaks) or are not reproducible (including the aggregation T_m at 48°C, which was detected in only one of three replicates).

As FT-IR spectra were acquired as a function of temperature, the succession of spectra can be analyzed by single value decomposition (SVD) [17]. This is an algebraic data reduction procedure which decomposes the succession of spectra into components with distinct temperature variation and the respective temperature profile. Importantly, the algorithm clusters the contribution of noise into separate components. The main components

Table 4.1 – Midpoint denaturation temperatures (T_m) for different secondary structure elements of ferredoxin. Values are extracted from the temperature profile of local spectra second derivative maxima or minima. Secondary structure assignment was performed according to their typical absorption wavenumbers [13].Temperatures are shown as average ± standard deviation of 2 to 4 independent measurements. Arrows indicate positive (↑) or negative (↓) second derivative peaks. Not coop.: transition is not cooperative.

	1st Temperature Ramp			**2nd Temperature Ramp**		
Assignment	**Band $\tilde{\nu}$ (cm^{-1})**	**pD 2.5**	**pD 12**	**Band $\tilde{\nu}$ (cm^{-1})**	**pD 2.5**	**pD 12**
		T_m (°C)			T_m (°C)	
β-sheets	↑ 1636	49 ± 5	47	↑ 1642		52 ± 1
	↓ 1629	**53 ± 3**		↑ 1640	27 + 78	
	↓ 1625	**59 ± 1**		↑ 1625		69 ± 7
				↓ 1625	**Not coop.**	
α-helices	**↓ 1660**	**44 ± 1**	**50±2**	↑ 1659		48 ± 0
	↑ 1653	47 ± 3	47	**↓ 1653**	**43 ± 7**	
Turns	↑ 1699	46 ± 1	49±1			
	↓ 1691	**47 ± 3**	**49±1**	↑ 1692		48
	↓ 1678	**43**	**47**			
	↑ 1673	58 ± 0	58±0			
Aggregation	↑ 1685	46 ± 2	48±1	**↓ 1685**		**48**
	↑ 1615	57 ± 3		↑ 1614	Not coop.	

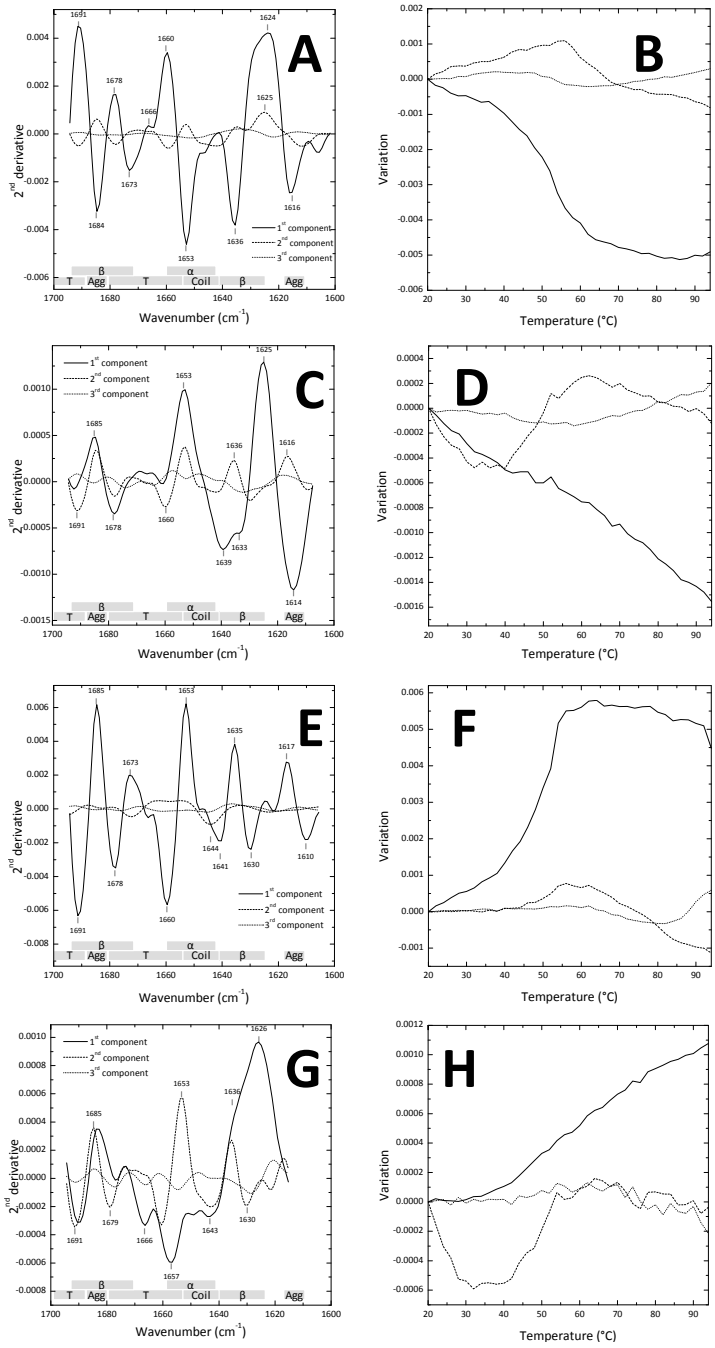


Figure 4.6 – Singular value decomposition of *A. ambivalens* ferredoxin FT-IR monitored thermal denaturation. (Left column) Component spectra computed from the temperature-dependent change in FT-IR spectra. (Right column) Temperature profiles of spectral components. (A-B) pD 2.5, 1st T-ramp; (C-D) pD 2.5, 2nd T-ramp; (E-F) pD 12, 1st T-ramp; (G-H) pH 12, 2nd T-ramp. Wavenumber secondary structure assignment is indicated. α, α-helices; β, β-sheets; T, turns; Agg, intermolecular β-sheets, aggregation fingerprint; Coil, random coil.

calculated after SVD analysis were compared with the characteristic spectra of particular types of secondary structure as well as with the temperature profiles. The results of this analysis are shown in Figure 4.6.

The structural features being represented by SVD components can be determined by examining local (second derivative) minima. In the original spectra, these represent peak zones which are associated with secondary structure elements. Overall, most temperature-induced secondary structure changes are captured by the first two components. Accordingly, we have simplified further analyses by considering only these components. The structural characterization of each component as well as the T_m values for their unfolding are listed in Table 4.2.

Inspection of Table 4.2 reveals that distinct conformational changes are involved in ferredoxin's thermal unfolding. At pD 2.5 α-helices, β-sheets and turns are involved in all unfolding events. α-helices unfold during holo-ferredoxin denaturation. The second thermal denaturation only involves β-sheet and turn structures and occurs at lower temperatures (44°C versus 51 and 60°C) and with lower cooperativity even though the structural elements involved are the same. This is in accordance with the stabilizing effect of the FeS clusters. Thermal unfolding at pD 12 is more complex, with more

Table 4.2 – Structural and stability characterization of SVD components. Not coop.: not cooperative. Aggregation: intermolecular β-sheets.

		pD 2.5		pD 12	
	Component	**T_m (°C)**	**Structure**	**T_m (°C)**	**Structure**
Ramp 1	1	51	α-helices β-sheets Aggregation	51	β-sheets Turns
	2	60	β-sheets Turns	54	α-helices Coil
Ramp 2	1	Not coop.	β-sheets Aggregation	60	α-helices Coil Turns
	2	44	β-sheets Turns	48	β-sheets Aggregation Turns

structural transitions occurring. Nevertheless, the secondary structures and transition temperatures (T_m) are similar for the first and second temperature ramps.

It is clear that some components occur in several conditions. This includes components which occur in the independent assays carried out at pD 2.5 and pD 12. For example, the second component of both thermal denaturations at pD 2.5 are identical among themselves and to component 1 from the first thermal denaturation at pD 12 (Figure 4.6). In fact, this component presents multiple second derivative minima and these occur at wavenumbers characteristic to all secondary structure types occurring in ferredoxin (1617 and 1685 cm^{-1}: intermolecular β-sheets; 1673 cm^{-1}: β-turns; 1653 cm^{-1}: α-helices; 1635 cm^{-1}: intramolecular β-sheets). The different T_m for this common component (44, 51 or 60°C) is then reporting for its pD-dependent stability and the overall conformation of ferredoxin. This common SVD component indicates that the same conformational changes occur in the thermal denaturation assays carried out at dissimilar conditions and may thus be suggested as a fingerprint of ferredoxin core structure.

By performing all pair wise component comparisons we have further identified pairs of components which are not identical but share local similarity in specific spectral regions (Figure 4.7). Although with lower identity than shown before for the fingerprint component, the remaining spectral components are represented partially in many of the experimental conditions assayed. This suggests that the thermal unfolding pathway is not significantly changed between the two pD values assayed and that discrete structural events are responsible for the different nature of the final "denatured" state at the two conditions: the molten globule at pD 2.5 and the featureless unfolded state at pD 12.

After establishing a structural rationale for the second SVD component at pD 2.5, one can gain further insight into ferredoxin thermal denaturation by

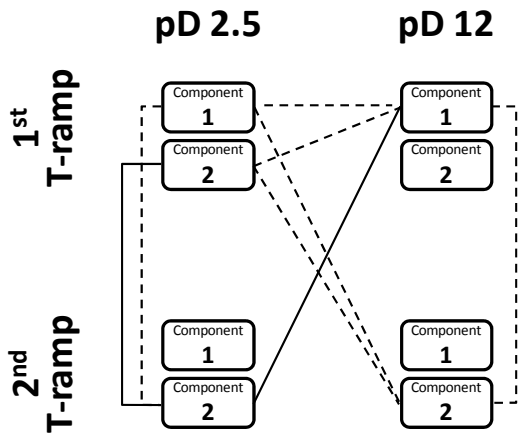


Figure 4.7 – Global SVD results analysis. Line style indicates spectral similarity (judged by second derivative peak position). Solid lines: component spectra are identical; Dashed lines: spectra are similar, with most peaks coinciding.

examining its temperature profile, which reports on the relative abundance of the ferredoxin core structure it describes. The temperature profile is not monotonic. Apart from the major transition reported in Table 4.2, at lower temperatures a compensating (*i. e.* opposite) conformational transition is observed (Figure 4.6), which resembles the cold unfolding transition in other proteins [19]. Protein cold unfolding is a fundamental biophysical phenomenon which is seldom observed because it frequently occurs at temperatures below the freezing point of water. The fingerprint SVD component indicates that the core ferredoxin conformation it represents undergoes both cold and a heat unfolding in experimentally accessible temperatures. The cold unfolding behaviour is also observed at pD 12 (second component, second temperature ramp, Figure 4.6). The component involved in this transition is similar to the one occurring at pD 2.5 but spectral differences indicate it represents a different ferredoxin conformation.

4.5. Discussion

In the current study we have used SVD decomposition of second

derivative FT-IR spectra to perform a structural characterization of *Acidianus ambivalens* ferredoxin thermal unfolding. This methodology enabled us to describe temperature-induced conformational changes which would not be detectable otherwise. Our analysis identified a fingerprint component spectrum describing a conformational involved in ferredoxin's thermal denaturation at acidic and basic pD. The formation of the ferredoxin molten globule has been described in detail [7] and provides the structural framework for the nature of the fingerprint SVD component. The molten globule is a semi-compact structure with tertiary structural similarity with the native state. The fingerptint component describes these conformational changes accurately, by being represented in the thermal denaturation of holo-ferredoxin (irrespectively of pD) and of molten globule ferredoxin but not in the denaturation of the protein previously unfolded at pD 12. Moreover, several SVD components describing thermal unfolding at pD 2.5 and pD 12 were found to be common or similar, showing conservation of unfolding pathways at the two pD conditions.

A previously unknown aspect in ferredoxin thermal denaturation is the involvement of a structural subset undergoind cold denaturation. This component – occurring at pD 2.5 and pD 12 – exhibits a complex, multi-stage temperature profile which is strikingly similar to the ones arising in cold denaturing proteins [19]. Cold unfolding is a general feature of protein folding. However, frequently it cannot be experimentally monitored because it occurs at temperatures below the freezing point of water. This phenomenon has been probed *in vitro* frequently using alcohols, which stabilize the cold unfolding transition. However, the perturbation introduced by the co-solvents may affect the protein's intrinsic stability properties and does not constitute a general strategy, suitable for all proteins. In this sense, the accessibility of the cold and heat unfolded states in ferredoxin – at least at acidic conditions – make it a unique model for describing protein folding

and unfolding thermodynamics in experimental conditions not involving external perturbations. The conformational plasticity of ferredoxin illustrated by the possibility of a conformational subset undergoing cold and heat unfolding at experimentally amenable conditions and in a reversible manner – the same component is observed in successive thermal denaturation assays – may thus be the biophysical basis for stable molten globule formation.

4.6. Conclusions

We have extended the biophysical characterization of *Acidianus ambivalens* ferredoxin folding properties by using an FT-IR approach. We have associated the previously identified aggregation step in ferredoxin unfolding with intramolecular β-sheet formation and have described a previously undetected low temperature α-helical and β-turn conformational transition. Additionally, we have carried out SVD analysis describing ferredoxin's thermal denaturation behavior in acidic and basic conditions. Our results suggest that the apo molten globule state of ferredoxin arising upon thermal denaturation at pD 2.5 may be a useful model for addressing cold protein unfolding, a phenomenon which cannot be studied in most proteins without the presence of additives like alcohols.

4.7. References

1. Eck RV & Dayhoff MO (1966) Evolution of the structure of ferredoxin based on living relics of primitive amino Acid sequences. Science 152, 363-366
2. Wachtershauser G (1988) Before enzymes and templates: theory of surface metabolism. Microbiol Rev 52, 452-484
3. Wittung-Stafshede P, Gomes CM & Teixeira M (2000) Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens. J Inorg Biochem 78, 35-41
4. Rocha R, Leal SS, Teixeira VH, Regalla M, Huber H, Baptista AM, Soares CM & Gomes CM (2006) Natural domain design: enhanced thermal stability of a zinc-lacking ferredoxin isoform shows that a hydrophobic core efficiently replaces the structural metal site. Biochemistry 45, 10376-10384
5. Frazao C, Aragao D, Coelho R, Leal SS, Gomes CM, Teixeira M & Carrondo MA (2008) Crystallographic analysis of the intact metal centres [3Fe-4S](1+/0) and [4Fe-4S](2+/1+) in a Zn(2+) -containing ferredoxin. FEBS Lett 582, 763-767
6. Moczygemba C, Guidry J, Jones KL, Gomes CM, Teixeira M & Wittung-Stafshede P (2001) High stability of a ferredoxin from the hyperthermophilic archaeon A. ambivalens: involvement of electrostatic interactions and cofactors. Protein Sci 10, 1539-1548

7. Leal SS & Gomes CM (2007) Studies of the molten globule state of ferredoxin: structural characterization and implications on protein folding and iron-sulfur center assembly. Proteins 68, 606-616
8. Todorovic S, Leal SS, Salgueiro CA, Zebger I, Hildebrandt P, Murgida DH & Gomes CM (2007) A spectroscopic study of the temperature induced modifications on ferredoxin folding and iron-sulfur moieties. Biochemistry 46, 10733-10738
9. Leal SS & Gomes CM (2008) On the relative contribution of ionic interactions over iron-sulfur clusters to ferredoxin stability. Biochim Biophys Acta 1784, 1596-1600
10. Armengaud J, Gaillard J, Forest E & Jouanneau Y (1995) Characterization of a 2[4Fe-4S] ferredoxin obtained by chemical insertion of the Fe-S clusters into the apoferredoxin II from Rhodobacter capsulatus. Eur J Biochem 231, 396-404
11. Moulis JM & Meyer J (1982) Characterization of the selenium-substituted 2 [4Fe-4Se] ferredoxin from Clostridium pasteurianum. Biochemistry 21, 4762-4771
12. Bertini I, Cowan JA, Del Bianco C, Luchinat C & Mansy SS (2003) Thermotoga maritima IscU. Structural characterization and dynamics of a new class of metallochaperone. J Mol Biol 331, 907-924
13. Barth A & Zscherp C (2002) What vibrations tell us about proteins. Q Rev Biophys 35, 369-430
14. Gomes CM, Faria A, Carita JC, Mendes J, Regalla M, Chicau P, Huber H, Stetter KO & Teixeira M (1998) Di-cluster, seven-iron ferredoxins from hyperthermophilic Sulfolobales. J Biol Inorg Chem 3, 499-507
15. Glasoe PK & Long FA (1960) Use of glass electrodes to measure acidities in deuterium oxide. J Phys Chem 64, 188-190
16. Pace CN, Shirley BA & Thomson JA (1990) Measuring the conformational stability of a protein. In Protein structure - a practical approach (Creighton TE, ed, pp. 311-330. IRL Press, Oxford.
17. Henry ER & Hofrichter J (1992) Singular value decomposition: application to analysis of experimental data. Methods Enzymol 210, 129
18. Barth A (2000) Fine-structure enhancement--assessment of a simple method to resolve overlapping bands in spectra. Spectrochim Acta A Mol Biomol Spectrosc 56, 1223-1232
19. Pastore A, Martin SR, Politou A, Kondapalli KC, Stemmler T & Temussi PA (2007) Unbiased cold denaturation: low- and high-temperature unfolding of yeast frataxin under physiological conditions. J Am Chem Soc 129, 5374-5375

5. Characterization of the folding and stability of a mesophilic rubredoxin

Rubredoxin was purified and supplied by Isabel Pacheco (ITQB).

5.1. Summary

Rubredoxins are good models for the biophysical study of metalloprotein folding due to their structural simplicity and amenability to spectroscopic analysis. We have characterized the folding and stability properties of the rubredoxin from *Desulfovibrio gigas* using a toolbox of conformation-sensitive spectroscopic techniques which allow describing secondary and tertiary structure changes complementarily: UV/visible absorption, circular dichroism (CD) and Fourier transform infrared spectroscopy (FT-IR). We show that the mononuclear iron site is thermodynamically and kinetically stable, withstanding high temperature and iron sulfur cluster disruption by EDTA stripping. We present a temperature-resolved model for the conformational changes occurring during rubredoxin unfolding derived from FT-IR spectra deconvolution: in the 20-60°C range α-helices and β-sheets unfold; above 60°C, aberrant turn structures and β-sheets buildup. The dual regime of rubredoxin unfolding with residual structure present at high temperature may account for this protein's high thermostability.

5.2. Introduction

Rubredoxins are the smallest and simplest iron-sulfur (FeS) proteins, containing a single iron atom coordinated by four cysteine sulfurs occurring in the consensus sequence CXXCG–()$_n$–CXXCG, where X indicates any amino acid residue. These proteins function as electron carriers in multiple metabolic processes, a process where the solvent exposed FeS cluster accommodates one electron. Due to their thermostability, [1] simplicity and incorporation of a simple metal cofactor, rubredoxins are useful models for the study of the molecular determinants of protein thermostability and the cross-talk between metal cofactors, protein folding, stabilization and function. For these purposes, two rubredoxins have been instrumental: the *Pyrococcus furiosus* (a hyperthermophilic archaeon) and the *Clostridium*

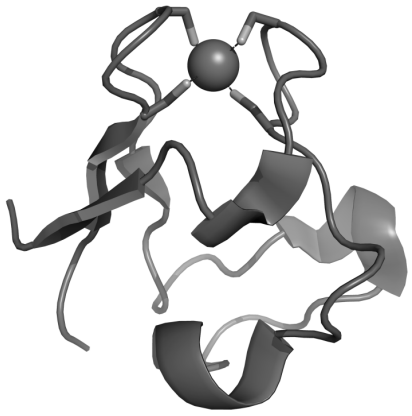


Figure 5.1 – Rubredoxin from *Desulfovibrio gigas*. The iron ion is shown as a sphere and the cysteine ligands as sticks. PDB: 2DSX.

pasteurianum (a mesophilic bacteria) proteins. The different stability of both proteins together with high resolution crystal structures has allowed establishing correlations between the molecular and stability features of both proteins.

In the current report we carried out a preliminary characterization of the folding and stability properties of the rubredoxin (Rd) from the mesophilic sulfate reducing bacteria *Desulfovibrio gigas* (Uniprot P00270-1). This is a 5.7 kDa protein whose structure is known from crystallography (Figure 5.1) [2] and NMR spectroscopy [3]. Also, theoretical [4] and spectroscopic [5] data are available from the literature, including the description of the stabilization of this protein by diglycerol phosphate [3]. However, a conformational characterization of the folding properties of this protein is lacking. To this purpose we have performed thermal denaturation of rubredoxin while monitoring its conformational status using absorption, circular dichroism and FT-IR spectroscopies. By combining these secondary structure sensitive techniques to spectral deconvolution we have observed high polypeptide and iron site thermostability and we propose a model describing the conformational changes occurring in the thermal unfolding of *D. gigas* rubredoxin.

5.3. Materials and methods

5.3.1. Rubredoxin

Desulfovibrio gigas rubredoxin was purified as previously described [2]. Protein concentration was determined spectrophotometrically using the

visible absorption extinction coefficients ε^{376nm} = 8450 $M^{-1}.cm^{-1}$ and ε^{493nm} = 6970 $M^{-1}.cm^{-1}$ [6]. Purity was verified by the absorption ratio between 377 and 278 nm equaling 0.45. Buffer was 50 mM potassium phosphate at pH 7.

5.3.2. UV/visible absorption

UV/visible absorption spectra were recorded using a Shimadzu UV-1700 spectrophotometer at room temperature.

5.3.3. Circular dichroism

CD measurements were recorded in a Jasco J-815 spectropolarimeter equipped with a Peltier-controlled thermostated cell support. 0.1 cm (far UV) and 1 cm (near UV/visible) pathlength cuvettes were used. Thermal denaturation experiments were carried out increasing the temperature from 25 to 95°C at a heating rate of 1°C/min. Every 5°C spectra were acquired. Thermal denaturation was assessed by the CD signal variations at local spectra maxima and minima. Midpoint denaturation temperatures (T_m) were obtained from sigmoidal fits to the normalized signal variation. Protein concentration was 0.1 mg/ml (17.6 µM, far UV) or 0.62 mg/ml (109 µM, near UV/visible). Temperature denaturations performed in the presence of EDTA (0.5 mM) were carried out at 0.1 mg/ml (17.6 µM, far UV) or 0.47 mg/ml (83 µM, near UV/visible) rubredoxin. Far UV CD spectra were accumulated 4 times at 200 nm/min scan rate and 1 s time response. Near UV/visible spectra were accumulated 5 times at 1000 nm/min scan rate and 0.5 s time response.

5.3.4. ATR FT-IR spectroscopy

Attenuated total reflectance Fourier transform infrared spectroscopy (ATR FT-IR) measurements were performed using a Bruker IFS 66/S spectrometer equipped with a nitrogen-cooled MCT detector using the thermostatized Harrick BioATR cell II. Protein was concentrated to 10 mg/ml and centrifuged at 12000 g before the temperature ramp to pellet any

aggregates forming while concentrating. FT-IR spectra in the amide I (1600-1700 cm^{-1}) and amide II (1500-1600 cm^{-1}) regions were recorded while temperature was increased from 20 to 94°C in a discontinuous fashion: spectra were accumulated during 1 minute (97 accumulations), temperature was raised 2°C during around 45 seconds and sample was equilibrated 15 seconds. Overall temperature change rate was 1.1°C/min. Spectral resolution was 4 cm^{-1}, scanner velocity was 20.0 kHz and aperture was 12 mm. Spectra were analyzed after vector normalization. Spectral variations at specific wavenumbers within the amide I band were assigned to secondary structure changes [7]. Thermal denaturation was assessed by the temperature dependence o absorption of absorption second derivative. Midpoint denaturation temperatures were determined from sigmoidal fits to the signal variation at local maxima or minima. Spectra deconvolution was performed as previously described [7]. Briefly, the amide I band was reconstituted as the sum of Gaussian curves centered at second derivative minima. Secondary structure content was estimated from gaussian curve integration and band assignment [7].

5.4. Results

5.4.1. Spectroscopic fingerprint of rubredoxin

Rubredoxin has multiple spectroscopic probes, many of those arising from the FeS cofactor. In the current study we have employed a toolbox composed of UV/visible absorption, circular dichroism and FT-IR spectroscopies. Far UV CD and FT-IR provide quantitative information about distinct secondary structure elements and are thus excellent methodologies to describe protein conformational changes and unfolding processes. By combining these techniques with the examination of the tertiary structure (near UV CD) and FeS cofactor integrity (visible absorption and visible CD) of *Desulfovibrio gigas* rubredoxin we provide a complete, multi-probe analysis of the stability

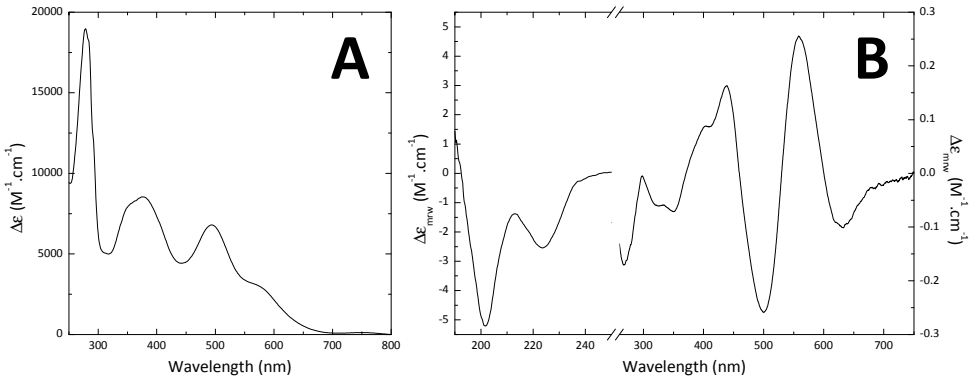


Figure 5.2 – Spectroscopic fingerprints of native *Desulfovibrio gigas* rubredoxin. (A) UV/Visible absorption spectrum. (B) Circular dichroism spectra. Left scale: far UV region (<250 nm). Right scale: near UV/visible region (>250 nm).

of this simple protein. The combination of visible absorption and visible CD spectroscopies to monitor the FeS cluster is especially powerful as the absorption signal relates to the FeS cluster integrity and the visible CD signal is rich in features and very sensitive to small changes in the cluster micro environment. In accordance, we have determined this protein's spectral fingerprints (Figure 5.2).

The UV-Visible absorption spectrum (Figure 5.2A) is typical for rubredoxin, with two bands in the visible region at 376 and 493 nm. Due to the chirality of the iron site, these bands are also CD active and give rise to multiple bands (Figure 5.2B), typical for the structural environment of the mononuclear iron site. The far UV CD spectrum shows two minima: one at 200 nm, characteristic of random coil-rich proteins and another at 223 nm, typical of β-sheet structures. This is in accordance to the native protein structure (Figure 5.1, [8]).

5.4.2. The rubredoxin fold is highly thermostable

To quantify the thermal stability of rubredoxin we have performed thermal denaturation assays monitoring protein unfolding by CD in the far UV, near UV and visible regions in the 25-95°C range (Figure 5.3). There

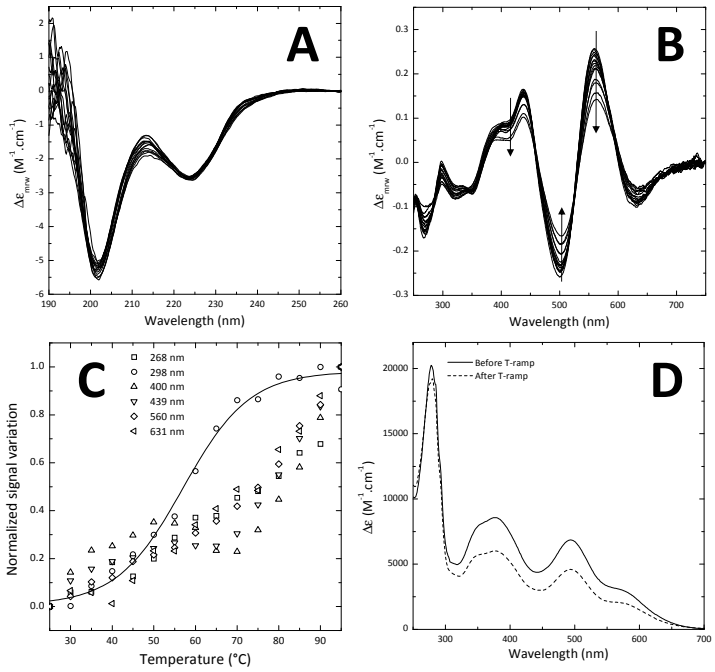


Figure 5.3 – Thermal denaturation of rubredoxin. (A) Far UV and (B) near UV/visible CD spectra recorded while increasing temperature from 25 to 95°C. Arrows indicate spectral changes with temperature. (C) Temperature-induced CD signal variation at specific wavelengths. A sigmoidal fit to the 298 nm signal change yields T_m = 57.0°C. (D) UV/visible absorption spectra before and after the near UV/visible-monitored thermal denaturation. The absorption change at 376 nm suggests a loss of 30% of the iron content.

were no temperature-induced changes in the far UV CD spectra (Figure 5.3A), indicating that secondary structure is kept. Tertiary contacts affecting the microenvironment around the single tryptophan side chain in this protein (CD signal at 298 nm) were disrupted at T_m = 57°C (Figure 5.3C). Nevertheless, some of the FeS clusters degraded, as noted by the decrease in the visible CD and absorption intensities (Figure 5.3BD). This process was irreversible even after extended incubation at room temperature. The protein is in fact highly thermostable: FeS cluster disintegration, as assessed by the variation in visible CD intensity (Figure 5.3C), is not complete at 95°C. The variation in absorption at 376 nm corresponds to a 30% FeS cluster disruption.

To examine the stability of the iron site and its contribution to the overall

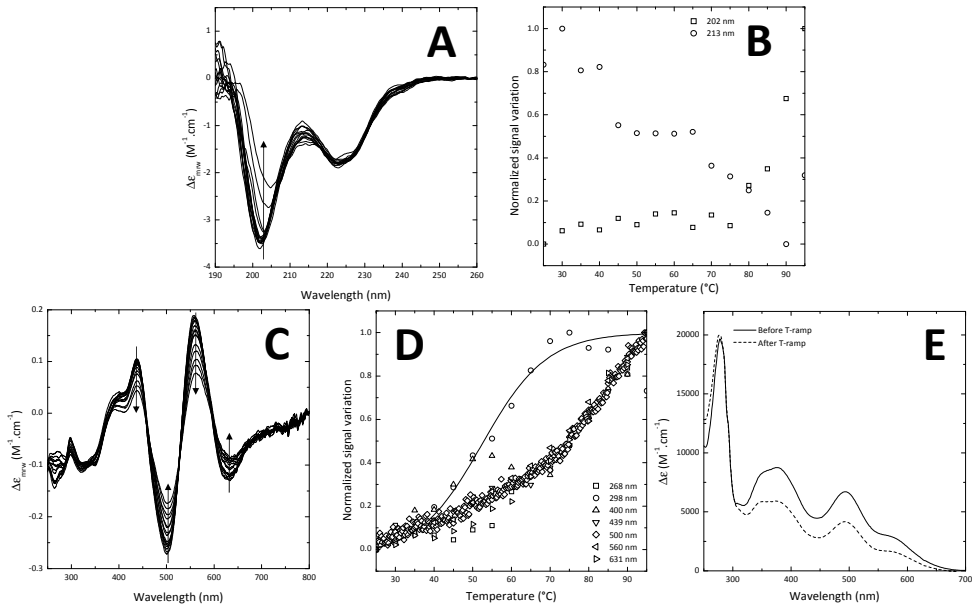


Figure 5.4 – Thermal denaturation of rubredoxin in the presence of EDTA. 0.5 mM EDTA was included in the buffer. (A) Far UV and (C) near UV/visible CD spectra recorded while increasing temperature from 25 to 95°C. Arrows indicate spectral changes with temperature. Temperature-induced CD signal variation at specific wavelengths in the (B) far UV and (D) near UV/visible regions. Sigmoidal fit to the 298 nm signal change yields T_m = 52.4°C. (E) UV/visible absorption spectra before and after the near UV/visible-monitored thermal denaturation. The absorption change at 376 nm suggests a loss of 33% of the iron content.

protein stability, we have carried out identical thermal denaturation experiments in the presence of 0.5 mM EDTA (Figure 5.4). This chelator should facilitate iron removal from rubredoxin, which constitutes an opportunity to assess its stabilizing role. EDTA destabilizes rubredoxin's secondary structure: the far UV CD bands at 202 and 213 nm are progressively lost at high temperature (>75°C, Figure 5.4AB). Since these bands are associated with random coil structures, this suggests full unfolding yielding highly dynamic and isotropic conformations. However, the destabilizing effect does not seem to be associated with destabilization of the iron site. There are no important changes in near UV/visible CD spectra apart from the slight destabilization of the 298 nm transition to T_m = 52.4°C (ΔT_m = 4.6°C, Figure 5.4CD). More importantly, FeS cluster loss estimation from visible absorption yields 33% (Figure 5.4E), close to the value obtained in the

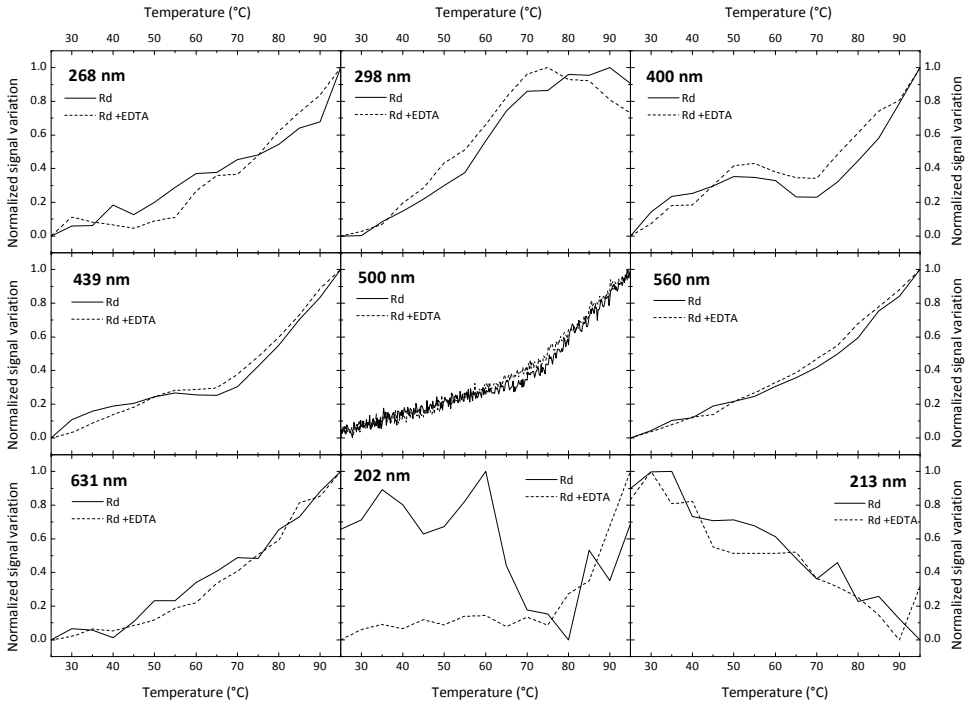


Figure 5.5 – Temperature stability of the rubredoxin iron site. Pairwise comparison of the visible CD signals obtained during thermal denaturation of rubredoxin in the presence and absence of EDTA. Data compiled from Figure 5.3C and Figure 5.4D.

absence of EDTA. The inefficiency of EDTA to affect the iron site during unfolding can be established by observing the thermal denaturation profiles of the visible CD signals in the presence and absence of EDTA (Figure 5.5). The pairwise analysis shows that in most cases the unfolding behavior is independent of EDTA. The conclusion is that under the conditions tested, the FeS cluster does not contribute significantly to protein stability.

5.4.3. Rubredoxin undergoes CD-undetectable unfolding

To have a more comprehensive view of rubredoxin's stability, we have monitored its thermal denaturation using FT-IR (Figure 5.6). Like CD, FT-IR can provide a quantitative estimation of protein secondary structure content. In comparison, FT-IR is more sensitive than CD to the conformation of

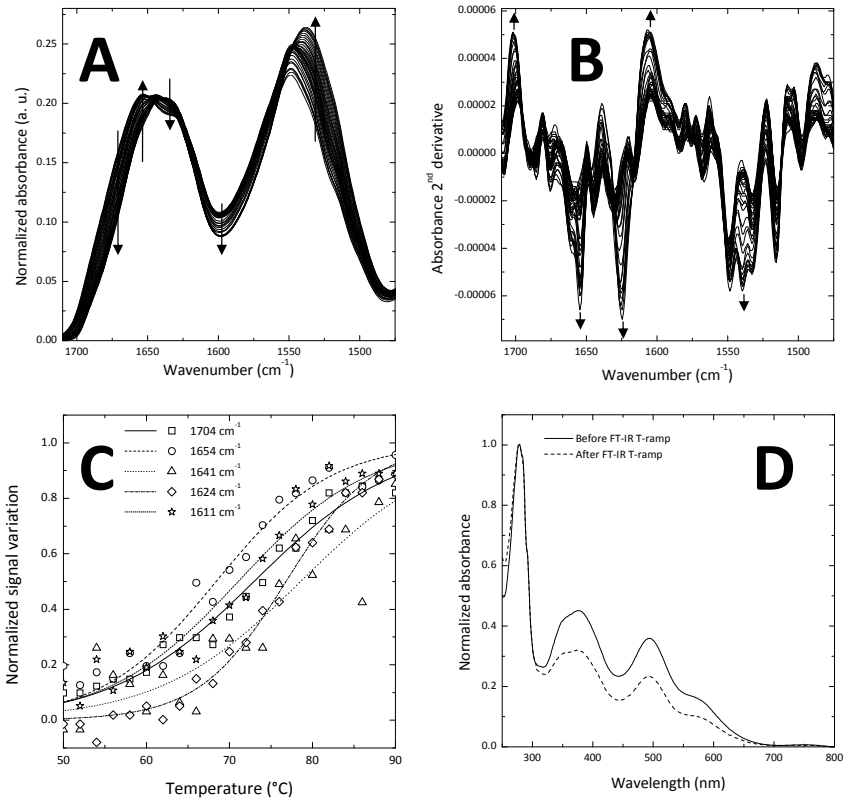


Figure 5.6 – FT-IR monitored thermal denaturation of rubredoxin. (A) FT-IR spectra in the amide I/amide II region. The trend of spectra recorded in the 20-94°C range is shown as arrows. (B) Second derivative of absorption spectra. (C) Temperature profile of second derivative peaks. (D) UV/visible absorption spectra obtained before and after thermal denaturation. Spectra were normalized so that $Abs^{278} = 1$. Quantification of iron loss from the absorption loss at 276 nm yields 29%.

proteins containing low helical content and β-sheets, which is the case of rubredoxin.

The FT-IR spectrum of native rubredoxin has a single maximum at 1639 cm^{-1} (Figure 5.6A, Figure 5.7AB), a wavelength characteristic of β-sheet containing proteins [7]. Upon incubation at high temperature, this single maximum gives rise to a peak at 1653 cm^{-1} and a shoulder at 1633 cm^{-1}. To extract information about the secondary structure types involved in the conformational transitions revealed by the FT-IR absorption spectra, we have considered the second derivative spectra (Figure 5.6B). This

mathematical manipulation enhances the detection of superimposed local peaks originating from different secondary structure contributions [7]. We can determine the order of unfolding of each secondary structure element by studying the thermal denaturation of the contributions of local maxima and minima (Figure 5.6C) and assigning them the secondary structure type with characteristic absorption at their wavenumber (Table 5.1). The most significant data are the ones originating from second derivative minima as they correspond to local maxima in the absorption spectra (*i. e.* reveal components with significant contribution to the overall spectra).

All secondary structure components undergo thermal transitions in the 68-79°C range, indicating that unfolding is multi state. The most relevant components (second derivative minima, Table 5.1 bold) correspond to α-helices (1654 cm^{-1}) and β-sheets (1624 cm^{-1}). These results then indicate that α-helical loss is the first event in the thermal denaturation of rubredoxin. This result is unlike the one obtained using CD, which revealed no significant secondary structural changes in the absence of EDTA. The iron loss percentage extracted from UV/visible absorption spectra is 29%, in the same range as the ones obtained during CD-monitored denaturation. This is an indication that the end conformation in all experiments is similar. In this way, the extra information provided by FT-IR is related to its enhanced

Table 5.1 – Thermal stability of secondary structure elements in rubredoxin as assessed by FT-IR. The temperature profiles of absorption at second derivative maxima and minima were fitted with sigmoidal curves to extract T_m values. Secondary structure assignments were preformed as in [7]. Second derivative minima (in bold) correspond to peak areas in the absorption spectra. Mix: in this zone there are contributions from β-sheet structures and tyrosine side chains. Aggregation: intermolecular β-sheets, frequently associated with aggregation.

$\tilde{\upsilon}$ (cm^{-1})	1704	**1654**	1641	**1624**	1611
Assignment	Mix	**α-helices**	Coil	**β-sheets**	Aggregation
T_m (°C)	73	**68**	79	**77**	71

sensitivity for this protein's conformation relatively to CD.

To have a complementary assessment of secondary structure changes occurring during rubredoxin denaturation, we have deconvoluted the FT-IR spectra. We have fitted the amide I band with sigmoidal curves centered at amide I maxima and minima and have assigned each band to the secondary structure element originating it (Table 5.2). By performing band fitting in a temperature-dependent manner, we can obtain a detailed temperature-resolved model of the conformational changes occurring along rubredoxin's thermal unfolding. To benchmark our deconvolution procedure, we have compared the structural deconvolution of the spectra at 24 and 94°C with the protein's crystal structure (Figure 5.7AB, Table 5.3). Overall, the deconvolution yields secondary structure contents compatible with the crystal structure, albeit with an overestimation of β-sheets and an underestimation of coiled regions.

From the temperature profile of individual components (Figure 5.7D) one can immediately identify the major structural rearrangements during thermal unfolding: increase of β-sheet (1643 cm^{-1}) and turns (1661 cm^{-1}). Other components have more irregular behaviors. This analysis indicates with accuracy the kind of secondary structure remodeling. For example, one can distinguish intermolecular from intramolecular β-sheets based in the absorption wavenumber. To have a global understanding or protein

Table 5.2 – Band assignment for the deconvolution of FT-IR spectra. β-sheets and amino acid side chains absorb in this region. Discarded for secondary structure quantification.

Wavenumber (cm^{-1})	Structure
1684	β-sheets
1661	Turns
1643	Coil
1639	Coil
1644	α-helices
1625	β-sheets
1603	mixed

Table 5.3 – Deconvolution of rubredoxin's FT-IR spectra. Comparison of the secondary structure content obtained from FT-IR spectra deconvolution with the one in the protein's crystal structure (2DSX [8]). n.a. not discriminated in pdbsum.

	Structural content (%)		
	Crystal structure	**24°C**	**94°C**
Helix	12	5	0
β-sheets	15	36	48
Turns	n.a.	1	31
Coil	73	57	21

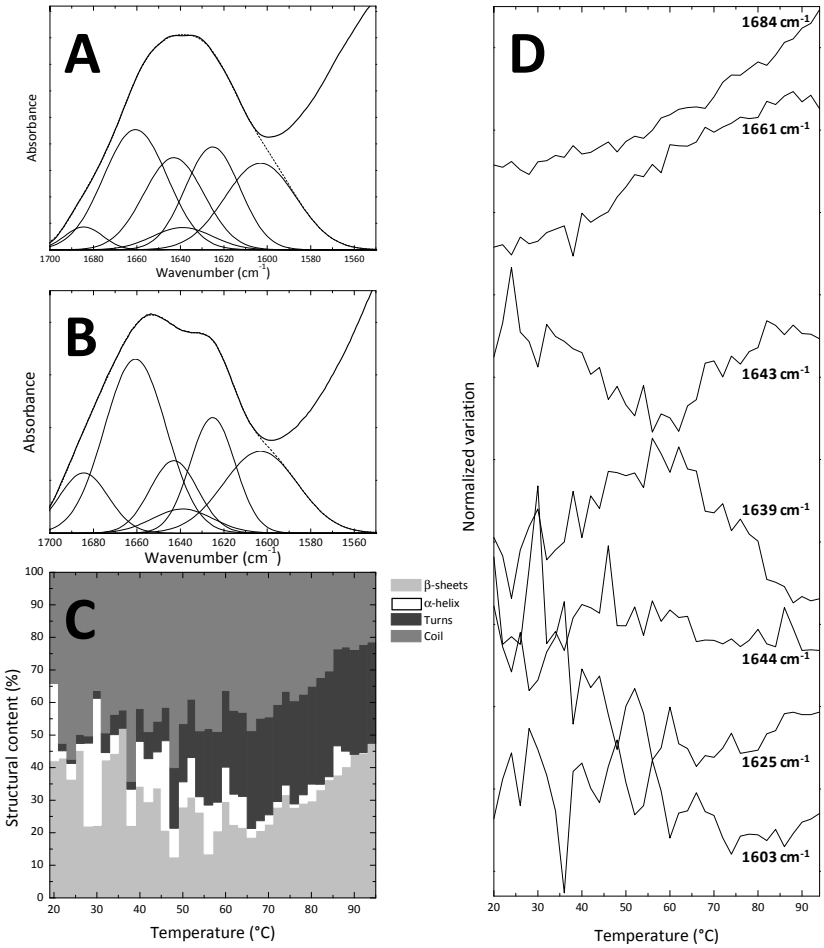


Figure 5.7 – Deconvolution of temperature-dependent FT-IR spectra. Spectra recorded at 24°C (A) and 94°C (B). The thick solid lines correspond to the spectral whose components were identified as gaussian bands (thin lines) whose sum (dashes) reconstitute the original spectra. By using the band assignment in Table 5.2, and the temperature profiles of each component band (D), the secondary structure changes with temperature were determined (C).

unfolding, one can sum the contributions arising from equivalent secondary structure types and study their temperature profiles (Figure 5.7C). This pictorial view confirms the results obtained with the second derivative approach. The first event in rubredoxin thermal denaturation is α-helical loss. At the same time, β-sheet content also decreases with concomitant buildup of random coil. Beginning at around 60°C, the trend changes with unfolding proceeding according to a distinct regime. Turns accumulate (~30%) as well as β-sheets, which return to native-like amounts at 94°C. These structures are formed from previously unstructured regions. The endpoint conformation is completely devoid of α-helices.

5.5. Discussion

In this work we have taken advantage of the combination of multiple spectroscopic techniques that probe similar molecular features to characterize the thermal unfolding of *D. gigas* rubredoxin. The dissimilar results originating from CD and FT-IR experiments occur because of the different sensitivities of each technique. CD spectroscopy is highly sensitive to α-helical elements because of their high far UV CD extinction coefficient [9]. The sensitivity for random coil structures or β-sheets is lower. In the case of rubredoxin this implicated that subtle conformational changes involving α-helical loss and gain in turns was undetectable. There is the possibility that the optimal (and different) protein concentrations for each technique influence results. However, aggregation was not observed in any condition. For this possibility, the FeS cluster percentage loss serves as an internal control. The similar iron loss in all assays suggests that the end conformation is similar, excluding the concentration artifact.

FT-IR deconvolution allowed a temperature-dependent description of the conformational changes composing the thermal unfolding of rubredoxin. This included the description of two regimes in the unfolding pathway: a lower temperature one (<60°C), wherein α-helices and β-sheets are lost; and

a high temperature one (>60°C), where turns and β-sheets accumulate. The additional information from CD-monitored thermal denaturation indicated that these are not global conformational transitions, as at 94°C far UV CD spectra report native-like structure and visible CD spectra indicate that the remaining FeS clusters are in an environment spectrally undistinguishable from that of the native protein. Nevertheless, a partial unfolding was simultaneous to FeS cluster degradation, an irreversible damage. Altogether, all data indicate a multi stage unfolding process.

Curiously, *Clostridium pasteurianum* rubredoxin seems to have similar unfolding properties to the *D. gigas* protein. It was reported that it undergoes major but reversible structural changes below FeS cluster dissociation temperatures, including lower compaction and local unfolding in the 50-70°C temperature range, the same where the *D. gigas* protein exhibits the unfolding regime change [10]. It has been shown that for the *Pyrococcus furiosus* rubredoxin FeS cluster release is rate determining in protein unfolding, unfolding only occurring after this step [11]. The possibility that the same process may occur in the *D. gigas* protein was left unaddressed. However, the fact that we have observed modest unfolding transitions and incomplete iron dissociation does not exclude this possibility.

The slight conformational destabilization of rubredoxin due to EDTA is likely the outcome of the enhanced iron release during unfolding or direct rubredoxin-EDTA binding. Thermal denaturation assays in the presence of increasing EDTA concentrations (to detect the concentration thresholds where the iron stripping effect would become significant) coupled to the examination of the effect of salt (a competitor for EDTA binding to rubredoxin) would clarify this issue. In any case, the fact that EDTA does not significantly increase iron release from holo-rubredoxin is an indication of high kinetic stability of bound iron.

Protein thermostability is the outcome of multiple evolutionary adaptations [12]. In the case if rubredoxin, a metal-containing protein, contributions can be attributed to the polypeptide itself and to the FeS cluster. Probably because of their small size and compact fold, rubredoxins are quite stable proteins. Regarding the FeS cluster, in the *C. pasteurianum* protein, the T_m for iron release in thermal ramp experiments is 69°C for Fe^{2+}-Rd and 83°C for Fe^{3+}-Rd [10]. The importance of stabilizing the FeS cluster site is highlighted in the *P. furiosus* protein, which contains an extended hydrogen bond network around this region as part of its stabilization strategies which enable cell growth at 100°C [13]. In other organisms similar adaptations exist. *C. pasteurianum* encodes a rubredoxin with an increased number of hydrogen bonds stabilizing the β-sheet [13].

5.6. Conclusions

In the present report we have described the thermal unfolding of *D. gigas* rubredoxin using UV/visible absorption, CD and FT-IR spectroscopies. Rubredoxin is a highly thermostable protein, retaining a significant amount of native-like secondary structure at high temperatures. The thermodynamic and kinetic stability of the iron mononuclear site seems to be significant as judged from the low thermal lability and insensitivity to EDTA. By deconvoluting FT-IR spectra obtained along thermal ramps, we have described unfolding as a two regime process, wherein at low temperatures (<60°C) α-helices and β-sheet unfold and at higher temperatures (>60°C) non-native β-sheets (48%) and turns (31%) buildup. The significant residual structure at high temperature may account for some of the protein's thermostability.

5.7. References

1. LeMaster DM, Tang J, Paredes DI & Hernandez G (2005) Contribution of the multi-turn segment in the reversible thermal stability of hyperthermophile rubredoxin: NMR thermal chemical exchange analysis of sequence hybrids. *Biophys Chem* **116**, 57-65

2. Bruschi M, Hatchikian C, Le Gall J, Moura JJ & Xavier AV (1976) Purification, characterization and biological activity of three forms of ferredoxin from the sulfate-reducing bacterium Desulfovibrio gigas. *Biochim Biophys Acta* **449**, 275-284
3. Lamosa P, Brennan L, Vis H, Turner DL & Santos H (2001) NMR structure of Desulfovibrio gigas rubredoxin: a model for studying protein stabilization by compatible solutes. *Extremophiles* **5**, 303-311
4. Micaelo NM, Victor BL & Soares CM (2008) Protein thermal stabilization by charged compatible solutes: Computational studies in rubredoxin from Desulfovibrio gigas. *Proteins* **72**, 580-588
5. Almeida RM, Pauleta SR, Moura I & Moura JJ (2009) Rubredoxin as a paramagnetic relaxation-inducing probe. *J Inorg Biochem* **103**, 1245-1253
6. Gomes CM (1999) *Oxygen utilization by prokaryotes*. PhD, Universidade Nova de Lisba, Oeiras.
7. Barth A & Zscherp C (2002) What vibrations tell us about proteins. *Q Rev Biophys* **35**, 369-430
8. Chen CJ, Lin YH, Huang YC & Liu MY (2006) Crystal structure of rubredoxin from Desulfovibrio gigas to ultra-high 0.68 A resolution. *Biochem Biophys Res Commun* **349**, 79-90
9. Kelly SM, Jess TJ & Price NC (2005) How to study proteins by circular dichroism. *Biochim Biophys Acta* **1751**, 119-139
10. Bonomi F, Fessas D, Iametti S, Kurtz DM, Jr. & Mazzini S (2000) Thermal stability of Clostridium pasteurianum rubredoxin: deconvoluting the contributions of the metal site and the protein. *Protein Sci* **9**, 2413-2426
11. Cavagnero S, Zhou ZH, Adams MW & Chan SI (1998) Unfolding mechanism of rubredoxin from Pyrococcus furiosus. *Biochemistry* **37**, 3377-3385
12. Vieille C & Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* **65**, 1-43
13. Bougault CM, Eidsness MK & Prestegard JH (2003) Hydrogen bonds in rubredoxins from mesophilic and hyperthermophilic organisms. *Biochemistry* **42**, 4357-4372

Part II – Metal ions and protein folding in the S100 family: Calcium, Zinc and Copper

6. Metal ions and protein folding

6.1. Metals in biology

Although being highly versatile macromolecules, proteins could not fulfill all their required functions relying only on the properties of the polypeptide chain. Cofactors are frequently involved in protein stabilization, catalysis, signal transduction and ligand binding. Transition metal ions work as superacids in enzyme active sites and can increase the electrophilicity or nucleophilicity of reacting species, increase the acidity of a reacting species and promote heterolysis [1]. More importantly, without cofactors polypeptides are not able to perform oxygen transport or proper electron transfer. In this case, biological systems have evolved to harness and control the unique properties of bioavailable metal ions [2-4]. Reciprocally, the polypeptide chain fine tunes the cofactor's properties according to functional requirements. Special relevance can be given to metal ion cofactors. More than 30% of all proteins incorporate metal ions in their polypeptide chains [5-6]. 47% of all enzymes incorporate metals, the majority at the respective active sites [1, 7]. Iron-containing compounds have been proposed to be primordial energy sources, instrumental in the development of the first cell,

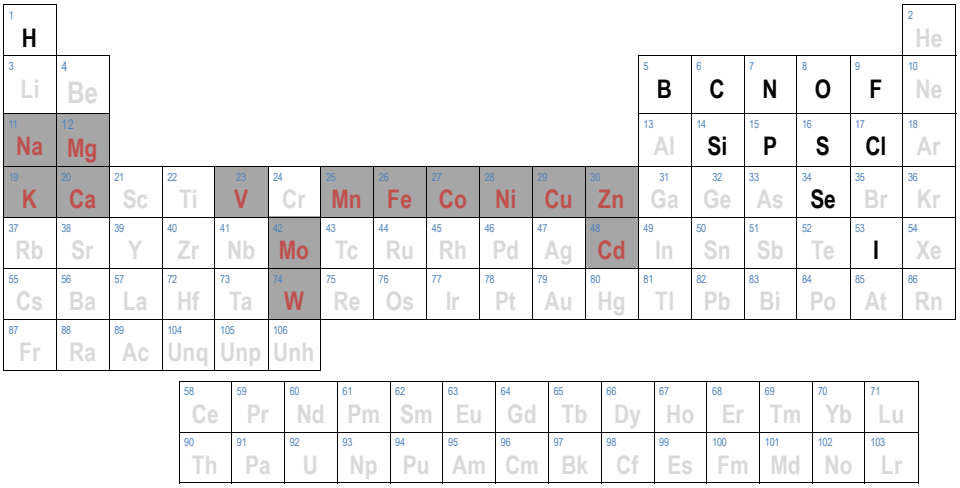


Figure 6.1- Bioinorganic periodic table. Metals with biological roles are shown in red; other biologically relevant elements are in black. From [10].

which then became the first protein cofactors upon amino acid ligand coordination [8-9].

Redox-inert transition metal ions are selected by proteins because of their structural stabilization properties. Mg^{2+} and Zn^{2+}, which are by far the most common metal cofactors, fulfill this role (Figure 6.2A). Distinct from Zn^{2+}, which is sometimes found permanently bound to proteins, Mg^{2+} is often bound to phosphate groups of substrates for electrostatic reasons and interacts with proteins only transiently. Comparatively, Ca^{2+} is less used. The redox-active metal ion most commonly found in redox centers is iron, followed by manganese, cobalt, molybdenum, copper and nickel (Figure 6.2B) [1]. Organic moieties sometimes associate with metal ions as a way of controlling their redox properties. This happens with iron and nickel, and most frequently with cobalt and molybdenum.

Wolynes and coworkers have estimated that around 10% of the interactions in native proteins are frustrated, meaning that they enclose a latent stabilization potential which is not used because of competing interactions. The frustrated interactions tend to cluster in the protein surface, being associated with protein-protein interaction [11-12] or ligand

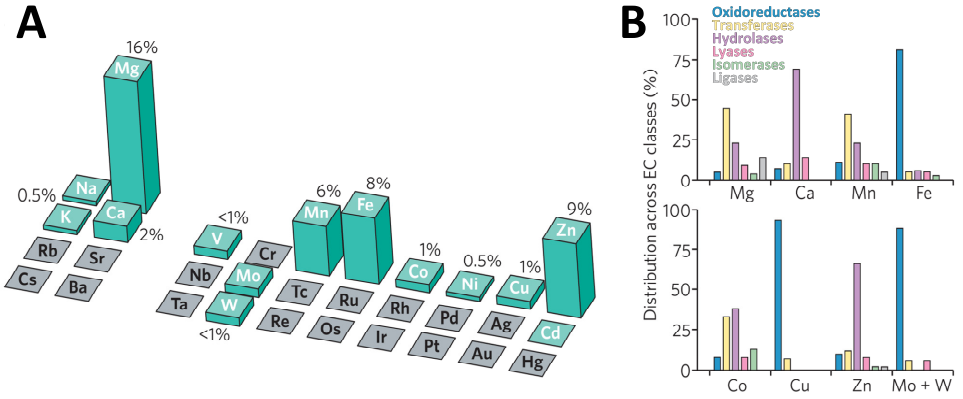


Figure 6.2- Metal ions used in catalysis. (A) Proportion of each metal used by all enzymes. Only one enzyme uses cadmium. Grey color indicates elements with no biological role. (B) Metal utilization among Enzyme Commission (EC) classes: oxidoreductases (blue), transferases (yellow), hydrolases (purple), lyases (pink) and ligases (grey). Counts are normalized so that the sum of the frequency for all classes equals 100% for each metal. From [7].

binding regions. Consequently, the underlying energetics is the driving force for protein structuring related to these specific associations.

The unique properties of metal ions make them indispensable for living organisms, which however keep metal ion homeostasis through a complex machinery which constantly adjusts the metal ion concentration and availability according to needs (Figure 6.3). This task is of the utmost importance as metal ions have promiscuous binding properties and can be engaged in uncontrolled redox reactions. These might lead to oxidation of cellular constituents with broad consequences, ranging from inhibition of essential enzyme activities, destruction of cellular compartmentalization to, ultimately, cell death. Sodium, magnesium and potassium are bulk metal ions, present in millimolar concentrations in the free state which keep the physiological electrolytic balance. Compartmentalization of each ion ensures an appropriate composition of the intracellular and extracellular fluids. The free concentration of all other metal ions is kept very low. For example, the total intracellular copper concentration in yeast is 70 μM. However, copper delivery to SOD1 relies on specialized machinery because free copper concentration is not high enough for spontaneous binding to occur despite the high binding affinity ($K_d \sim 10^{-15}$ M) [13]. Similarly, cellular free zinc concentrations are on the order of 10^{-13} – 10^{-9} M [14] and free calcium concentrations around 10^{-7} M [15-16].

Zinc or calcium are intracellular second messengers [17-18], that are released from cellular storage upon extracellular stimulus. Dedicated

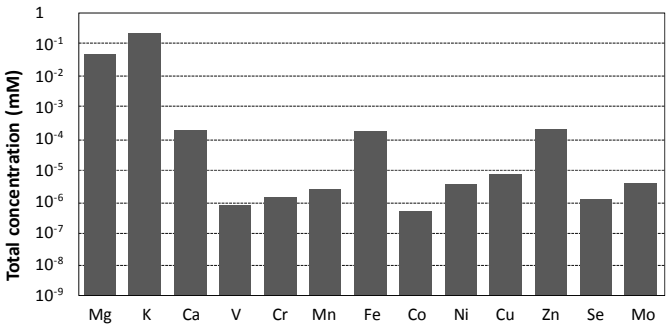


Figure 6.3 - *Escherichia coli* metallome. Total concentration of selected metal ions in cells grown in rich culture medium. Cells accumulate the appropriate amount of each metal ion. From [26].

transport and delivery systems exist to insert metal ions in their specific proteins – the metallochaperones [19]. For copper, there exist the Atx1-like chaperones, the CCS chaperones for superoxide dismutase and the copper chaperones for cytochrome *c* oxidase [19]. Iron is chaperoned by ferrochelatase or several scaffold proteins for iron-sulfur cluster assembly [20-21]. There is no zinc chaperone system. However, practically all cellular zinc is kept bound to metallothioneins, reduced glutathione and other proteins. Metallothioneins do also regulate the homeostasis of other physiological (*e. g.* copper, selenium) and xenobiotic (such as cadmium, mercury, silver and arsenic) metals which bind cysteine thiol groups (~30% of its amino acidic residues) [22]. Other redundant and non-specific metal ion homeostatic mechanisms like ion exchange or pumping to specific cellular compartments couple the control mechanisms of several metal ions. The outcome is that cellular metal homeostasis is a complex mechanism wherein the responses to the fluctuation of the concentration of one metal ion are coupled to some extent to the response regarding other metal ions (Figure 6.4) [23].

In the absence of metal delivery systems, metal incorporation is determined by free metal concentration [24] and the stability of the metal complex, which is generalized by the Irving-Williams series [25]:

$$Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$$

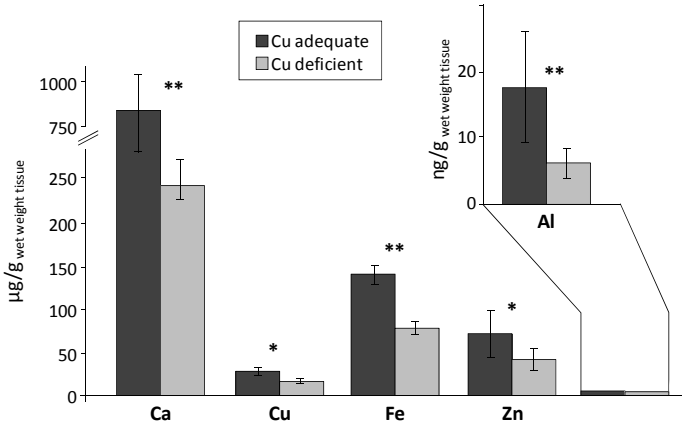


Figure 6.4 - Intertwined homeostasis between different metal ions. Metal ion amounts in CD-1 mice fed for 3 months with a Cu-adequate or a Cu-deficient diet. All metal ion levels are significantly affected by the dietary copper deficiency. * p-value < 0.05, ** p-value <0.01 versus control. From [23].

6.2. Shaping the conformational space: metal ions

The role of metal ions in protein folding is complex. The first step to understand this relationship is addressing how metals bind to proteins in the first place. Several binding modes have been described (Figure 6.5). Some metalloproteins are able to fold in the absence of the metal ion, which is necessary for functional conformational readjustments or catalytical purposes, like most calcium-binding proteins (*e. g.* calmodulin and S100 proteins). In other cases, the newly synthesized apo protein is not able to acquire its native fold. Zinc finger domains can only fold to the native structure in the presence of zinc. On the other hand, metal ions may bind to the nascent polypeptide chain as it emerges from the ribosome. In fact, heme binds to the unfolded state of cytochrome b_{562} [27-28] and copper remains bound after unfolding of azurin [29-30] and the Cu_A domain of cytochrome *c* oxidase [31].

Tightly binding metals can be permanently bound and be crucial for the acquisition of the native protein structure. Loosely bound metals can bind and rebind to proteins according to the metal concentration gradient. In this case, binding may be associated with conformational changes, dependent of the ligand bridging role of the metal. The concept of metal-induced folding or metal-induced conformational changes is associated with the mechanism of metal ion insertion into proteins. When they are synthesized in the ribosome, proteins emerge to the cytosol as unfolded or partially folded species. In some cases, in the absence of the metal ligand the newly-synthesized protein rests in a native-like fold (*e. g.* S100 proteins); metal ligation then determines conformational changes, being reversible in signaling-related proteins (Figure 6.5, blue pathway). In other situations metal ion incorporation is required for folding (*e. g.* Zinc fingers). It may occur either cotranslationally (Figure 6.5, green pathway) or posttranslationally (Figure 6.5, red pathway). Metal ion selection is determined either by the Irving-Williams binding

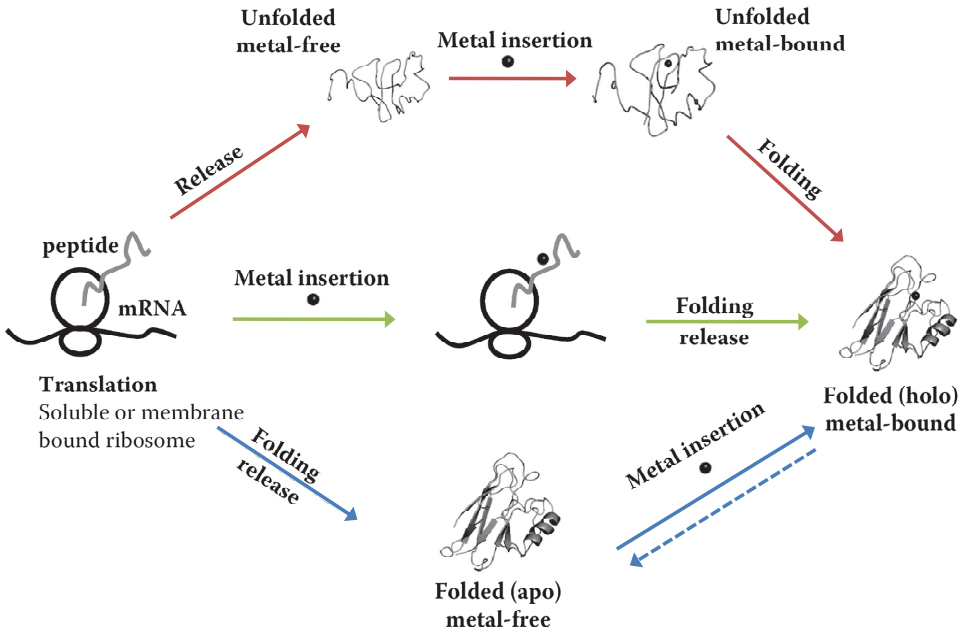


Figure 6.5 – Routes for the folding of metalloproteins. From [32].

stability series and metal ion concentration or by the action of metallochaperones.

In the next sections we will address the issue of the interplay of metal ions in protein folding by focusing in the most relevant metal ions – iron, zinc, copper and calcium – and illustrating this subject with representative examples.

6.2.1. Iron

Iron is the most abundant transition metal ion in the human body. Iron is a very versatile element due to its two stable oxidation states (Fe^{2+}/Fe^{3+}) which allow it to mediate redox reactions in a broad potential range and its several spin states which catalyze reactions involving molecular oxygen [10]. Iron is incorporated in proteins in the form of a heme, iron-sulfur clusters, mixed metal centers, di-iron or mononuclear centers. One of the best studied models of the interplay between iron and protein folding is cytochrome *c*, an electron transfer protein operating in the mitochondrial respiratory chain. In

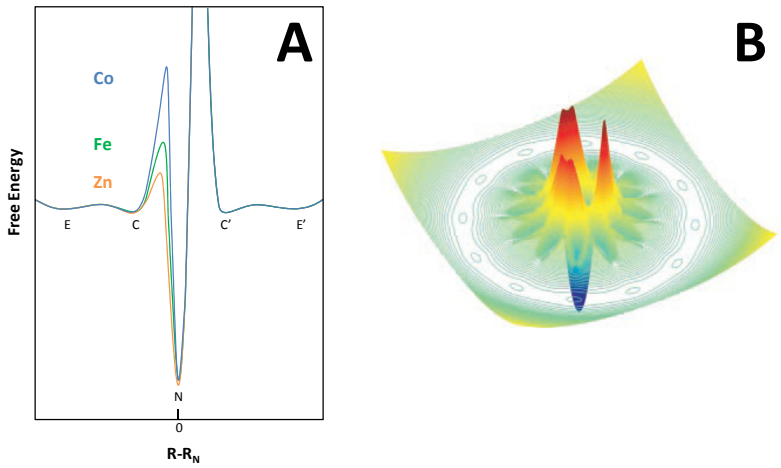


Figure 6.6 - Cytochrome *c* folding energy landscape. (A) Two dimensional cross-section of the folding energy profile for the cobalt, iron and zinc-substituted proteins. The folding reaction coordinate ($R\text{-}R_N$) represents the size (R) difference to the native protein size (R_N). Besides the native state, one can identify energy basins for expanded (E, E') and collapsed (C, C') intermediates. The metal substitution has a significant influence on the folding energy barrier. (B) Three-dimensional representation of the folding energy landscape for the native (iron-bound) protein. From [5, 34-35].

cytochrome *c* the iron is inserted in a heme group which is covalently bound to the polypeptide. The coordination is octahedral, the equatorial positions corresponding to the heme ligands and the two axial sites being filled by amino acid side chains. Cytochrome *c* is a highly stable protein, with a T_m close to 100°C [33]. The metal ion in the heme can be replaced *in vitro* with important consequences to protein folding (Figure 6.6). Substitution of Fe^{3+} by Zn^{2+} lowers the folding energy barrier but, of course, renders the protein redox inactive. Substitution by Co^{2+} raises the activation energy.

Heme removal is associated to secondary and tertiary structure loss [36]. Hence, heme binding is thought be a prerequisite for cytochrome *c* folding. However, this conception has been challenged by the recent observation of spontaneous folding of the *Aquifex aeolicus* apo cytochrome c_{555} [37]. An intermediate situation occurs with apo cytochrome c_{551} from *Pseudomonas aeruginosa*. This protein unfolds reversibly and acquires a compact, though marginally stable, conformation [38].

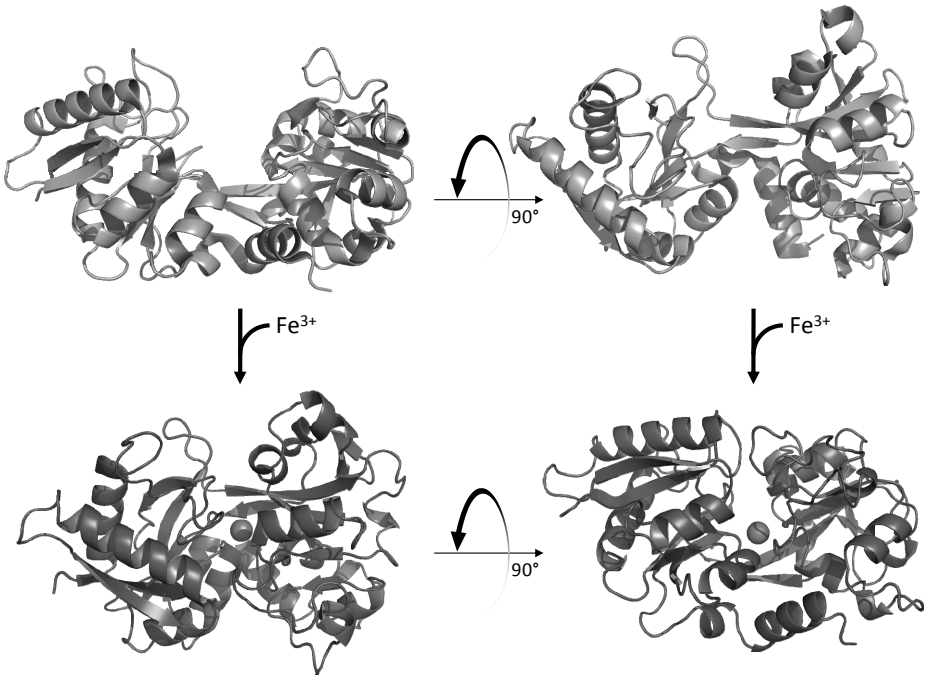


Figure 6.7 – Fe^{3+}-dependent conformational changes in human transferrin. The apo form (1BP5, top) adopts an open conformation. When Fe^{2+} (shown as a sphere) binds, the two protein domains close around the inter-domain hinge enclosing the metal ion in the protein core (1A8F, bottom).

The transport of iron in the blood is carried out by transferrin. Transferrin binds Fe^{3+} with an unusually high affinity (K_d = 10^{-22} M) [39]. Upon Fe^{3+} binding, the two transferrin domains undergo a conformational change which compacts the protein structure and buries the Fe^{3+} ion in the protein (Figure 6.7). Iron binding greatly stabilizes the protein and alters the unfolding pathway from an apparent three state ($T_{m,1}$ = 58°C, $T_{m,2}$ = 69°C) to a simple two state process (T_m = 87°C) [40]. Mutation of the Asp63 ligand decreases the Fe^{3+} binding affinity by approximately 5-6 orders of magnitude, with a concomitant destabilizing of ΔT_m ~12°C for the holo protein and ~3°C for the apo protein [41].

6.2.2. Zinc

Zinc is essential for life and the second most abundant transition metal ion in living organisms. Zinc is a redox-inactive metal ion due to its filled *d*

shell, $3d^{10}4s^2$, valence electronic configuration. However, it binds to proteins very tightly and fulfills structural roles in proteins. The most common coordination number for zinc is four. In proteins, three ligands are enough for ensuring tight binding, the fourth one being water. A recent survey has revealed that zinc-binding proteins are a significant part of all genomes (4-10%) [42]. The percentage of zinc binding proteins correlates with growth temperature: organisms thriving at higher temperatures make more extensive use of zinc-binding proteins. Underlying this preference may be the stabilizing role of zinc.

The most notable example of zinc-associated folding occurs in proteins containing zinc finger domains. Zinc fingers are the most abundant structural domain in the human proteome (present in 3% of all proteins) and the most abundant zinc binding motif [43-44]. [44] The classical zinc finger domain has a ββα fold, where the Zn^{2+} ion is coordinated by two cysteine ligands near the turn of the antiparallel β-sheet, and two histidine ligands in the C-terminal part of the α-helix (Figure 6.8A). Although the domain contains a small hydrophobic core [45], the apo form is unstructured. The structure can only fold and be brought into close proximity upon zinc binding (Figure 6.8B). Moreover, the zinc-bound domain is stabilized and becomes

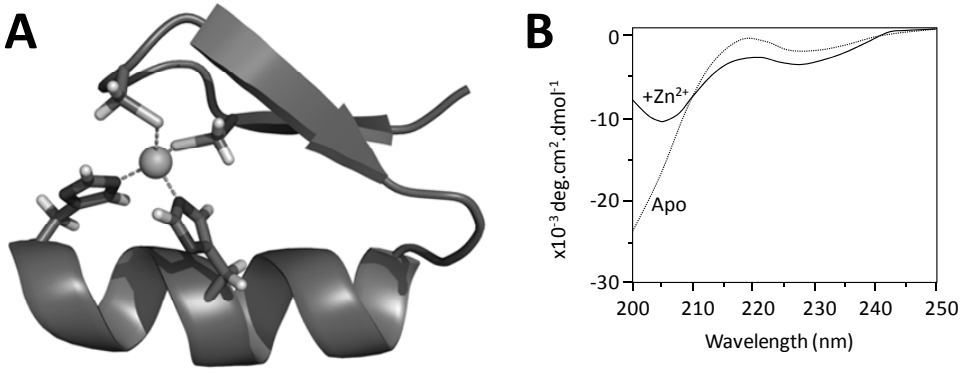


Figure 6.8 – Folding of a zinc finger domain. (A) Structure of a metallated Cys_2His_2 type zinc finger. (B) Far UC CD spectra of *Xenopus* transcription factor IIIA in the absence and presence of zinc, indicating the secondary structure occurring upon metal binding. From [46].

competent in DNA binding [46]. The zinc binding affinity to zinc finger domains is high (K_d = 10^{-11} to 10^{-9} M) [47], rendering them sensitive to transient Zn^{2+} concentration rises. In this manner, zinc finger domains couple zinc concentration variation to transcriptional responses and interaction with multiple proteins [48].

Zinc is one of the cofactors of the cytosolic Cu,Zn-superoxide dismutase (SOD1), the main protection against the highly oxidant superoxide radical. It does not take part in catalysis but is determinant for the acquisition of the protein's native β sheet-rich fold [49] and high thermostability. Yeast alcohol dehydrogenase is a tetrameric enzyme incorporating two Zn^{2+} ions in each subunit. One ion sits in the catalytic site; the other one has a structural role. Apart from activity, the incorporation of Zn^{2+} has been described as essential for proper folding, as it confers higher pH stability, higher stability towards chemical denaturation [50] and higher folding cooperativity [51-52]. The catalytical zinc ion is thought to have a more important effect in stability [50].

6.2.3. Copper

Azurin contains a single copper ion, which can be substituted *in vitro* by zinc. The protein structure is identical in the apo and metallated (Cu/Zn) states [53-54]. However, metal ions have a significant effect in the protein's thermodynamic stability. The stability of the apo protein is 6.9 kcal/mol. The binding of either Cu^{+} or Zn^{2+} raise the stability to ~9.5 kcal/mol, and Cu^{2+} to 12.4 kcal/mol [55]. Both calcium and zinc bind the native as well as the unfolded state of azurin. This has important consequences in the folding kinetics: Cu^{2+} binding before folding greatly accelerates the kinetics (milliseconds) relatively to metal binding after apo protein folding (minutes to hours) [55]. This has been proposed as a mechanism to ensure fast folding and selective binding of copper over zinc (which binds ~17000 fold less tightly).

Ceruloplasmin is the major copper protein in the blood and an essential regulator of iron homeostasis. It binds six copper ions, three forming a trinuclear cluster and all others in mononuclear sites. Holo ceruloplasmin is more stable than the respective apo protein (ΔT_m = 15 to 20°C). The metal load also influences the denaturation kinetics. In physiological-like conditions the apo protein unfolds is <2 days to a molten globule-like state [56], while the holo protein forms an intermediate state lacking one copper ion which is active and remains stable for more than two weeks, that was proposed to be relevant *in vivo* [57]. Ceruloplasmin also binds Ca^{2+} and Na^{+}. Na^{+} is thought to restrict the mobility of surface loops creating an entry point for substrate to the active site [58].

6.2.4. Calcium

Like magnesium, calcium is an inert and labile ligand. Evolution has exploited this property and rendered calcium an important second messenger, which is released from intracellular stores upon stimulation and binds reversibly to several proteins producing a transient signal. After release, calcium is rapidly sequestered by calcium-binding proteins or actively pumped outside the cell or to storage organelles. This is equivalent to say that calcium works as a trigger, activating the proteins it binds to. There are a variety of calcium binding motifs but the most abundant is the helix-loop-helix EF-hand (Figure 7.3), present in calmodulin, troponin C, the S100 proteins, among others [10]. Upon binding to EF-hands, calcium induces a large conformational change, shifting the position of one of the EF-hand helices and exposing a hydrophobic surface for protein-protein interactions (Figure 6.9A). Moreover, the electrostatic neutralization at the binding site upon cation binding is frequently responsible for protein stabilization (Figure 6.9B).

One of the most well studied EF-hand protein is calmodulin, a ubiquitous protein found in all eukaryotes which regulates enzyme activities,

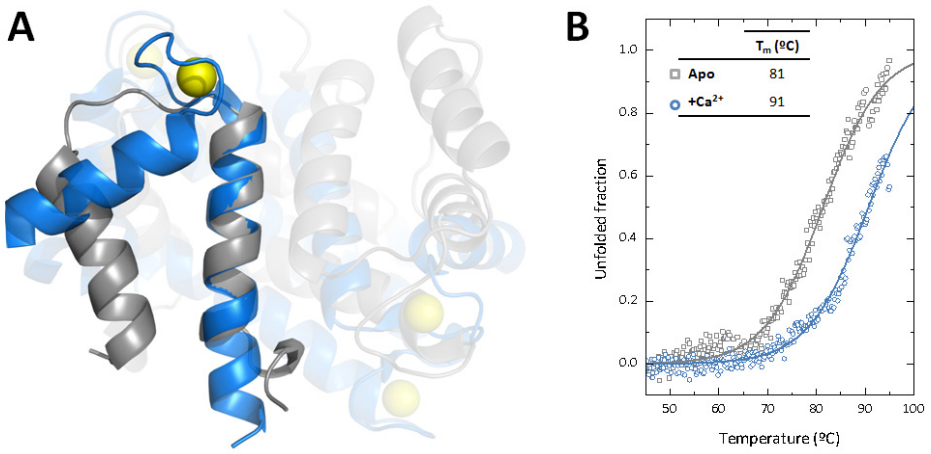


Figure 6.9 – Calcium-dependent conformation and stability of EF-hands. (A) Overlay of the apo (1K9P, grey) and Ca^{2+}-bound (1K9K, blue) S100A6 structures. The C-terminal EF-hand is highlighted to show the helix displacement occurring upon the binding of calcium (shown as yellow spheres). The inter-helical surface is hydrophobic and upon exposure becomes the docking site for downstream-acting proteins. (B) Thermal stability of S100A6 monitored by far UV CD in the apo and Ca^{2+}-bound states (0.1 mg/ml protein, 5mM KPi pH 7, 98 µM Ca^{2+}, λ=222 nm). Upon calcium binding the protein is significantly stabilized (ΔT_m = 10°C) (unpublished observations).

neurotransmitter release, cell proliferation and DNA repair. Most of these activities require interaction with other proteins in a calcium-dependent manner. Many of these proteins do not bind calcium themselves; they rely on calmodulin as a calcium sensor and signal transducer. Upon calcium binding calmodulin binds to nitric oxide synthase (NOS) and activates the electron transfer pathway necessary for catalysis [59], binds to microtubule-associated proteins, favoring microtubule depolymerization [60], interacts with kinases responsible for neurite outgrowth [61].

The archetypical EF-hand is the one in parvalbumin [62], where it was identified. Calcium has important relevance in the folding of this protein [10]. In the calcium-bound state the protein rests in a compact globular state, with the EF-hand in the open state (*e. g.* Figure 6.9A). Calcium release induces a fast (~60 ps) conformational rearrangement resulting in a decrease in the native helical content (>20%, [63]) and enhanced residue solvent accessibility. The apo state has been described as natively disordered [64-

65]. A subsequent structural analysis revealed a calcium-induced hydrophobic core repacking [66]. Regarding the protein dynamics, the calcium-bound protein becomes highly resistant to trypsin proteolytic cleavage [67].

Bovine α-lactalbumin is another EF-hand protein. It regulates lactose synthesis though calcium-dependent conformational changes. Calcium increases its folding rate by three orders of magnitude [68-70] and couples refolding to the generation of the native disulfide bridge [71], which suggests a calcium regulation of lactose synthesis through controlling α-lactalbumin folding. A more careful examination of the process revealed that calcium accelerates the folding of α-lactalbumin by decreasing the energy barrier between the molten globule and the transition states, by decreasing the difference of entropy between the two states [70]. The metal binding properties of α-lactalbumin are promiscuous and it also binds other metal ions apart from Ca^{2+}. Metal binding translates into specific protein stability (Figure 6.10).

Unlike hen egg white lysozyme, equine lysozyme binds calcium [72-73]. At acidic pH and in the absence of calcium the α and β domains unfold at distinct temperatures. The calcium-binding β domain is stabilized by the metal ligand. [74]. At neutral pH the protein unfolds in one cooperative

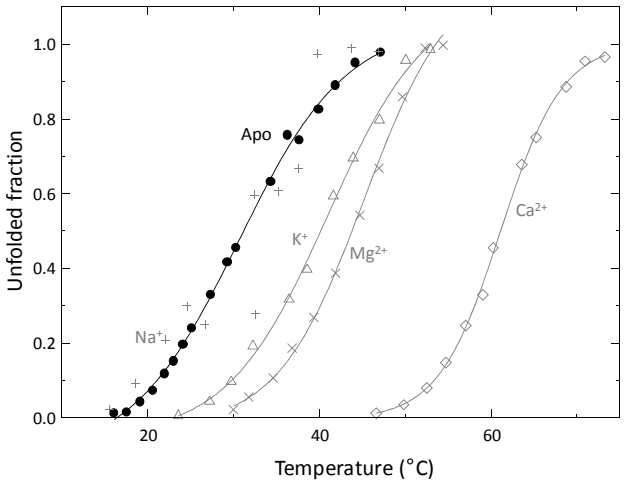


Figure 6.10 – Bovine α-lactalbumin thermal stability as a function of metal substitution. Thermal denaturation curves from intrinsic tryptophan fluorescence. Binding of Ca^{2+} to the apo protein brings about a rise in the T_m of about 30°C. Na^+ has no effect. K^+ and Mg^{2+} have intermediate effects. From [10].

transition in a calcium dependent melting temperature.

6.3. Expanding the conformational space: amyloidogenesis

Despite the solid evidence underlying it, Anfinsen's thermodynamic hypothesis of protein folding has been shown not to be completely valid. Namely, the postulate that the native state is unique and represents the most thermodynamically stable conformation is challenged by amyloid-forming proteins. These structures were first identified depositing as β-sheet rich aggregates at specific organ locations in association with disease (Figure 6.11A). These structures, could be selectively stained with a starch (amylose) selective dye, Congo Red (Figure 6.15), and were accordingly named as amyloid. Amyloid is a fibrillar conformation characterized by a cross-β structural motif emerging from X-ray fibril diffraction patterns [75-77]. Even though very distinct proteins are now known to form amyloid, they all exhibit a remarkably similar amyloid structure (*e. g.* Figure 6.11BC) wherein β-strands running nearly perpendicular to the fibril axis are stacked through hydrogen bond contacts. Structural polymorphism occurs at the detailed β-strand alignment and the stacking stabilizing interactions [78].

The amyloid state represents an "alternative" protein conformation and can be considered an extreme case of protein conformational change. Indeed, for all amyloid-forming proteins a "classic" native state exists – even though for many proteins it is a natively disordered one – which, under determined conditions can interconvert to amyloid (Table 6.3). In the cellular context, amyloid fibrils form spontaneously, although as a "rare" event associated with ageing and diseases including Alzheimer's, type 2 diabetes, prion diseases, Parkinson's, systemic amyloidosis and Huntigton's. Nevertheless, amyloid fibrils are remarkably stable, resisting proteolytic cleavage, chaperone-mediated disaggregation and, in some cases, even physicochemical denaturation. Thus, they must populate the bottom of the folding funnel. The "rare" amyloid formation in physiological context points

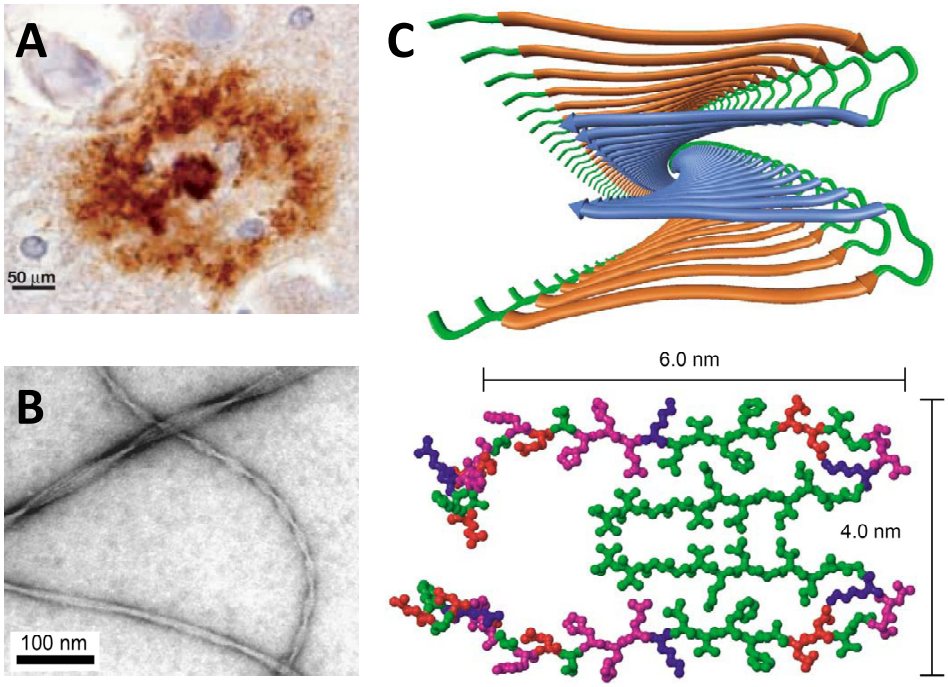


Figure 6.11 - Structure of Aβ amyloid structures. (A) Brain senile plaque stained with anti-Aβ antibodies. (B) Transmission electron microscopy image of Aβ fibrils. (C) Structural model of Aβ fibril ultrastructure. The fibril is formed by β-turn-forming monomers stacked along the fibril axis. The slight lateral offset between adjacent monomers introduces a twist in the fibril. From [78-79].

to amyloidogenesis being represented by a folding funnel concurrent with the native one, from which it is separated by a large energy barrier [80] (Figure 6.12). In this framework, the native state must be considered a kinetically trapped state which can only convert to amyloid under specific conditions.

Amyloid forming proteins were once viewed as a rare occurrence in biology, frequently associated with disease and not always reproduced in physiologically accessible conditions. Dobson and coworkers have shown otherwise. In fact, they have described amyloid formation by the SH3 domain of p85α subunit of bovine phosphatidylinositol 3-kinase, a artificial protein construct not involved in disease [82]. This discovery was the first evidence that, given the appropriate conditions, all proteins could, given adequate

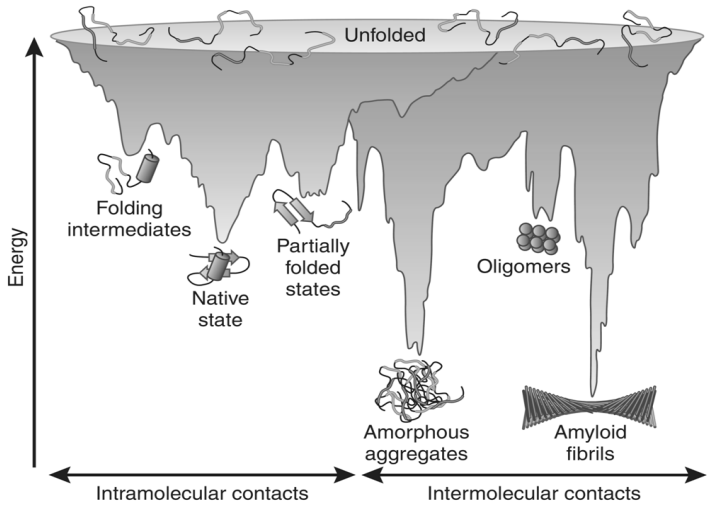


Figure 6.12 – Expanded protein folding funnel, considering aggregation processes. The "classical" folding scenario consists in the acquisition of the native conformation from the unfolded state, through a series of intermediates (blue funnel). Partially folded intermediates may expose hydrophobic patches which, especially at high protein concentration like the crowded cellular milieu, may drive aggregation of these misfolded proteins. Growing evidence supports the proposal that these partially folded species may facilitate the structural conversion and polymerization of amyloid structures of the oligomeric or fibrillar type (purple funnel). From [81].

conditions, adopt the amyloid conformation [82-83]. Much data has strengthened this view, of which one of the strongest ones is that all amyloids display a set of common structural features independently of the protein identity, pointing that at least some amyloidogenesis factors should be associated with the protein backbone rather than with amino acid side chains. In this scenario, the initial association of amyloid with a few proteins involved in disease results from the fact that amyloidogenic conditions for these proteins are in the physiologically-accessible range.

The infrequent amyloid formation *in vivo* as opposed to the postulate that the amyloid conformation should be accessible to all proteins suggests that evolution must have designed protein sequences for structure, fast folding kinetics, dynamics, stability, activity as well as for minimizing the amyloid-formation propensity. The latter must be a strong driving force because, owing to their high temperature [84] and proteolysis [85] resistance,

Table 6.1 - Functional amyloids. From [87-88].

Amyloid protein	Source	Function
Chaplin	*Streptomyces coelicor*	Aerial hyphae formation
Hydrophobin EAS	*Neurospora crassa*	Aerial hyphae formation
Curlin	*Escherichia coli*	Biofilm formation, cell adhesion, invasion
eRF3	Saccharomyces *cerevisiae*	Polyamine regulation ([PSI$^+$] phenotype)
Rnq1	Saccharomyces *cerevisiae*	Cytoprotection ([RNQ$^+$] phenotype)
HET-s	*Podospora anserina*	Heterokaryon formation
Ure2p	Saccharomyces *cerevisiae*	Promote the uptake of poor nitrogen sources ([URE3$^+$])
CPEB	*Aplysia californica*	Long-term memory
Silk moth chorion protein	*Lepidoptera*	Protection
Fish egg	*Austrofundulus limnaeus*	Resistance to dehydration
Eggshell chorion proteins	*Bombyx mori*	Protection
Pmel17	Mammals	Regulation of melanin biosynthesis
Fibrin	Mammals	Activation of haemostatic factors
Spidroin	*Nephila edulis*	Spider web silk

amyloids are not easily removed from biological systems and tend to accumulate. Insulin fibrils have strength and stiffness comparable to those of steel and silk, respectively. The amyloidogenic process is avoided by sequence design (*i. e.* by avoiding amyloid-prone sequence stretches), fast folding (minimizing aggregation), folding cooperativity (by disfavoring extensive conformational changes) and several cellular surveillance systems.

Thorough investigations have further revealed that some amyloids do have physiological importance. Some organisms make use of their mechanical performance or conformational templating properties to fulfill specific requirements (Table 6.1). Remarkable examples of these functional amyloids are spider silk – whose strength is higher than steel – and the yeast prions eRF3 and Ure2p – whose amyloidogenesis is associated with alteration in their native nucleic acid processing properties and constitutes an epigenetic, protein-only phenotypic inheritance mechanism [86].

6.3.1. Mechanisms of amyloid formation

Being a polymerization process, amyloid formation must be considered in the general framework of protein oligomerization and aggregation (Figure

6.13). Because of the gradual conformational changes occurring *en route* to the amyloid state, the intermediate oligomeric species also contain some amyloid features. One interesting aspect is that many of the proteins undergoing aggregation and deposition in amyloid diseases are globular and stable under physiological conditions. Nevertheless, the amyloid state in which they are found in disease situations is a highly stabilized state with a near universal structure, distinct from the native conformation, which is thus expected to occupy a disjoint space in the folding energy landscape. Then, the intriguing question is which conformational changes take place in this

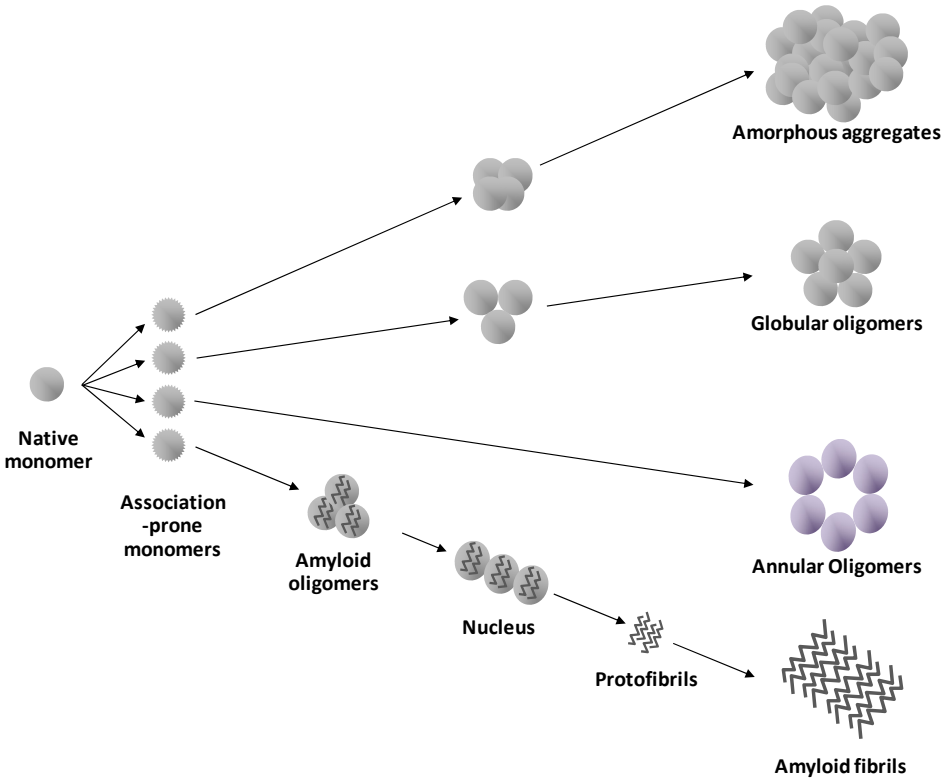


Figure 6.13 - Schematic representation of protein self-association processes. The same or distinct proteins can self-associate into oligomer structures with different structural features. Structured oligomers may actually constitute the native protein structure or be associated with signaling processes. The structural conversion from globular to amyloid conformations occurs through a series of intermediates. Once a nucleus is formed, fibrillization promptly proceeds by polymerization extending from these sites. Actual oligomerization processes are more complex and different structures and interconversions can occur. From [79, 97].

extensive structural conversion and how can the energetic barrier between the native and amyloid funnels be overcome. Chiti and Dobson have suggested that partial protein unfolding as a result of mutations, pH fluctuations, metal binding or imbalancement in cell protein quality control systems are responsible for the appearance and/or not elimination of partially unfolded/misfolded intermediates which may be at the beginning of the amyloidogenesis pathway [89]. More importantly, this hypothesis does not require extensive, energy costly, global conformational changes, as the misfolded state is still native-like. Acidic organelles like the endosomes and the lysosome have been associated with pH-induced generation of amyloid-like precursors of $A\beta_{1-42}$ [90], the prion protein [91] and transthyretin [92] but the opposite, clearance effect was observed for α-synuclein [93]. In the case of natively unfolded proteins similar mechanisms may apply. The structures of such intermediates have been pursued by theoretical and experimental methods [94-96].

The mechanisms of amyloid formation have been extensively studied *in vitro* for a variety of proteins, with the main focus on those involved in pathology. Five different types of mechanisms have been proposed (reviewed in [98]):

1. Subsequent monomer addition mechanism, where fibril elongation occurs by successive addition of monomers to a preformed fibril;
2. Reversible association mechanism, which considers the interaction between all kinds of species (instead of just the monomer and fibril) and aggregation reversibility;
3. Prion aggregation mechanisms, which explain prion infectivity by considering that the conformational transition of the native to the amyloid-prone conformational is thermodynamically unfavorable but aggregation is highly favorable and drives amyloid polymerization;

4. "Occam's razor"/minimalistic model (or Finke–Watzky 2-step model), which simplifies the analysis by considering separate contributions from nucleation and polymerization and representing all aggregation steps as simple kinetic steps;
5. Quantitative structure activity relationship models, which rationalize amyloid aggregation rates as a function of the physico-chemical properties of the aggregating protein.

Being an aggregation process, amyloid formation is based on three fundamental physical processes:

1. A critical concentration below which the frequency of molecular collisions is not enough for promoting aggregation;
2. A critical nucleus, which forms at the beginning of amyloid formation and seeds further fibril elongation;
3. Lag phase, which corresponds to a delay in apparent polymerization due to conformational changes and nucleus formation time.

The critical nucleus is defined as the least thermodynamically stable species in solution, which is the oligomer of minimal size capable of initiating further growth [99]. Fragmentation of existing aggregates also supplied nucleus-like structures [100-101]. The addition of preformed amyloid structures enhances the polymerization rate of the native protein. This is called amyloid seeding. The analysis of amyloidogenesis kinetics with mechanistic models allows extracting quantitative data which describe the fibrillization process: thermodynamic parameters, rate constants, cooperativity and nucleus size, among others [98].

For Aβ, a model was proposed according to which fibrils elongate by addition of monomeric Aβ at the edges following a "dock-and-lock" mechanism [96]. According to this mechanism, monomers are added to the growing fibril in several steps: first the monomer interacts with the fibril tip in a readily reversible manner; next conformational changes convert the

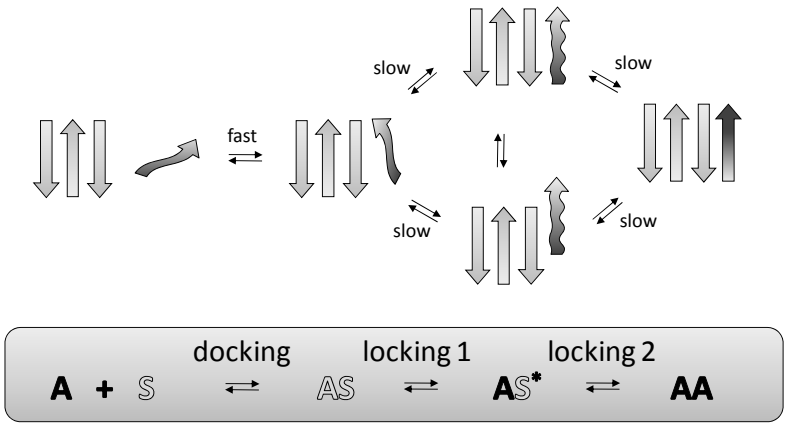


Figure 6.14 - Dock and lock mechanism of amyloid fibril growth. A: preformed amyloid; S: soluble precursor. According to this model, amyloid propagation from a preformed nucleus involves an initial fast step where the soluble precursor exchanges between the free state and a docked state where it is associated to the nucleus in a native-like conformation. The docked precursor then undergoes a conformational transition to an amyloid-like state which is prone to converting to the amyloid state by nucleus templating. From [96].

peptide to the amyloid conformation and the monomer is integrated into the fibril (Figure 6.14).

6.3.2. Detecting amyloids

Amyloid formation is a structural modification accompanied by changes in spectroscopic and other physical properties. The process can be probed and amyloid species detected using the same methodologies used for any type of protein aggregation (Table 6.2). The classical methods for unambiguously determining if a protein aggregate is of the amyloid type are the characteristic apple-green birefringence under a polarized light microscope after staining with Congo red – the standard method in

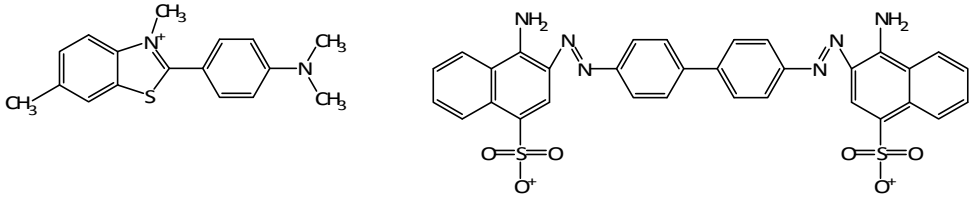


Figure 6.15 – Structure of thioflavin T and Congo red.

Table 6.2 - Physical methods to analyze protein aggregation. "Concentration dependent" techniques are those which can provide *in situ* analysis if aggregation conditions are within detection limits. [98].

Method	Information	*In situ*/*ex situ*	Kinetics ?
Absorbance	Direct	Concentration dependent	Yes
Atomic force microscopy	Direct	*In situ*	Yes
Calorimetry	Direct	*In situ*	Yes
Circular dichroism	Direct	Concentration dependent	Yes
Dyes	Indirect	Concentration dependent	Yes
Electron microscopy	Direct	*Ex situ*	No
EPR	Indirect	Concentration dependent	Yes
Flow birefringence	Direct	*In situ*	Yes
Fluorescence spectroscopy / intrinsic fluorophore	Direct	Concentration dependent	Yes
Fluorescence spectroscopy / extrinsic fluorophore	Indirect	Concentration dependent	Yes
FT-IR	Direct	Concentration dependent	Yes
Light scattering	Direct	*In situ*	Yes
Mass spectrometry	Direct	*Ex situ*	No
NMR	Direct	Concentration dependent	Yes
Quartz crystal oscillation	Direct	Concentration dependent	Yes
Turbidity	Direct	*In situ*	Yes
Viscosity	Direct	*In situ*	Yes
X-ray diffraction	Direct	Concentration dependent	No

immunohistochemistry – the characteristic thioflavin T fluorescence enhancement at 182 nm (Figure 6.15) or the cross-β fibril diffraction pattern.

The insoluble, heterogeneous, high molecular weight nature of amyloid fibrils poses great challenges to conventional structural characterization by X-ray fibril diffraction or NMR [102]. However, recent breakthroughs in fibril X-ray diffraction [102-104], solid-state NMR, neutron diffraction, microscopy and structural modeling have provided some evidence for the structure of amyloid fibrils. Still, only limited information is available about the structural conversion of natively folded proteins to amyloid structures. This is because despite the similarity of the amyloid fibrils formed by different proteins, the formation mechanism seems to be unique for each protein and the identity and structure of the intermediate conformational states is inherently

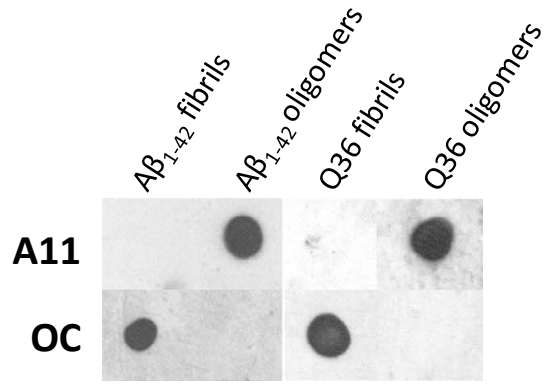


Figure 6.16 – Conformation-sensitive antibodies. Dot blot analysis of $A\beta_{1-42}$ and and a 36-mer polyglutamine peptide in the fibrillar or oligomeric forms detected with the fibril-specific OC antibody or the amyloid oligomer-specific A11 antibody. From [106].

heterogeneous, transient and masked by the multimolecular polymerization mechanism. To tackle these problems, Kayed and co-workers have raised conformation-dependent antibodies (reviewed in [105]), which detect the universal structure features of amyloid-like structures [106-107]: fibril structures (OC antibody) or amyloid oligomeric structures (A11 antibody) (Figure 6.16). These antibodies have revealed that the same structural features are observed on oligomeric or amyloid structures formed by very distinct proteins. Conformation-sensitive antibodies are privileged tools for identifying specific conformations in mixtures, the role of each type of structure in disease and for structural characterization of conformers in the amyloidogenesis pathway.

6.4. Metal ions and amyloid formation in disease

Protein aggregation in general and amyloid fibrillization in particular are very specific processes, involving orchestrated conformational transitions. Consequently, proteinaceous deposits associated with disease are found at specific organism locations and involving well defined proteins. Protein aggregation increases with aging and this can change the proteostasis network and enhance further aggregation [108] resulting in diseases (Table 6.3).

Protein fibrillization can be induced *in vitro* by mild destabilizing

Table 6.3 – Representative human amyloid diseases. Including the information about the disease related polypeptide. From [109].

Disease	Protein/Peptide	Native structure
Neurodegenerative diseases		
Alzheimer's disease	Aβ peptide	Natively unfolded
Spongiform encephelopathies	Prion protein	Natively unfolded & α-helical
Parkinson's disease	α-synuclein	Natively unfolded
Dementia with Lewy bodies	α-synuclein	Natively unfolded
Frontoremporal dementia with Parkinsonism	Tau	Natively unfolded
Amyotrophic lateral sclerosis	Superoxide dismutase 1	Immunoglobulin-like
Huntington's disease	Huntingtin, polyQ expansion	Mostly natively unfolded
Spinocerebellar ataxias	Ataxins with polyQ expansion	All β
Nonneuropathic systemic amyloidosis		
AL amyloidosis	Ig light chain or fragments	Immunoglobulin-like
AA amyloidosis	Serum amyloid protein A	All α
Senile systemic amyloidosis	Wt transthyretin	All β
Familial amyloidotic polyneuropathy	Transthyretin mutants	All β
Hemodialysis-related amyloidosis	β_2-microglobulin	Immunoglobulin-like
Apolipoprotein amyloidosis	Apolipoprotein fragments	Natively unfolded/unknown
Lysozyme amyloidosis	Lysozyme mutants	α+β
Nonneuropathic localized diseases		
Type II diabetes	Islet amyloid polypeptide	Natively unfolded
Atrial amyloidosis	Atrial natriuretic factor	Natively unfolded
Injection-localized amyloidosis	Insulin	All α
Corneal amyloidosis	Lactoferrin	α+β
Cataract	γ-crystallins	All- β

conditions like low pH, detergents or elevated temperature. However, the complex conformational changes involved and the concomitant protein fold perturbation have challenged our understanding of the mechanism responsible for *in vivo* amyloid formation. For some proteins, mechanisms for the amyloid conversion from the native conformation have been put forward [110]: intrinsic propensity for acquiring the amyloid conformation, which increases with age (*e. g.* α-synuclein in Parkinson's disease [111] or transthyretin in familial amyloid polyneuropathy [112]), increased protein

expression (*e. g.* α-synuclein gene locus triplication [113-114], hemodialysis-associated β_2-microglobulin accumulation [115], reduced insulin clearance from injection sites [116]), amyloid-favoring point mutations (including glutamine extension in polyQ amyloidosis), abnormal post-translational protein processing (*e*. g. tau hyperphosphorylation in Alzheimer's disease), altered proteolytic processing of the precursor protein (*e. g.* Aβ peptides in Alzheimer's disease) and environmental hazards capable of inducing conformational changes in the precursor protein (*e. g.* oxidative agents, heavy metals, pesticides). However, for the majority of amyloid diseases the causative reasons for amyloid aggregation are not known and most cases are sporadic.

In amyloid diseases protein misfolding is coupled to cellular degeneration, revealing that amyloid species are toxic. Several toxicity mechanisms have been uncovered: tissue integrity and function destruction by bulk invasion [110, 117], membrane destruction [118-119], induction of apoptosis by interacting with specific receptors [120], inhibition of neuronal calcium currents [121], impaired autophagosome maturation [122], autophagy dysfunction [123], oxidative stress induced by metal ion coordination by amyloid fibrils [124-128], exhaustion of cellular chaperoning capacity [129-130] and proteasome inhibition [131]. Systems biology approaches analyze the multiple perturbations associated with amyloid diseases as an imbalancement in the cellular proteostasis network [123]. In this sense, the accumulation of the toxic amyloid species can be understood as perturbing complex cellular systems and originating the several unexpected toxicity mechanisms associated with disease.

Metal ions are central players in amyloid formation. High affinity metal binding sites have been described in many amyloid-forming proteins, including β2-microglobulin [132-134], Aβ [135-138], α-synuclein [139-141], superoxide dismutase [49] and mammalian prion protein [142]. Metal ion

coordination may influence the conformational conversion of the native to the amyloid state or promote toxic reactivity. Iron and copper catalyze the generation of reactive oxygen species by cycling between their two accessible oxidation states. This can happen either in the free state or when bound to proteins. In either case, oxidative modifications can destabilize the native protein fold and contribute to amyloid formation. Even in the absence of any reactivity, metal ions can induce conformational changes which can determine the amyloid conversion. In the following sections we will illustrate the relevant role of metal ions in physiological amyloidogenic processes involved in disease focusing on neurodegenerative amyloid diseases.

6.4.1. Metal ions and Alzheimer's disease

Alzheimer's disease (AD) is the most common form of dementia. The major neuropathological lesions in AD patients are two kinds of amyloid deposits in the brain: amyloid-β (Aβ) peptide-containing neuritic plaques [143] outside neurons in limbic structures and association neocortex [144] and intracellular neurofibrillary tangles composed of hyperphosphorylated, polyubiquitinated tau protein [145] (Figure 6.17). Aβ and tau deposition seem to be driven by independent mechanisms as tau deposits occur in several other neurodegenerative diseases where no Aβ neuritic plaques are observed. This suggests that tau deposition is downstream from Aβ plaque formation but is an essential process in degeneration [146].

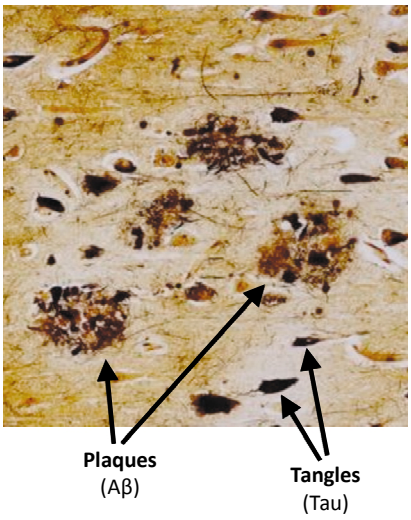


Figure 6.17 - Aβ plaques and tau neurofibrillary tangles. From [147].

Aβ is produced by the proteolytic processing of Amyloid Precursor Protein (APP, Figure 6.18), an integral membrane

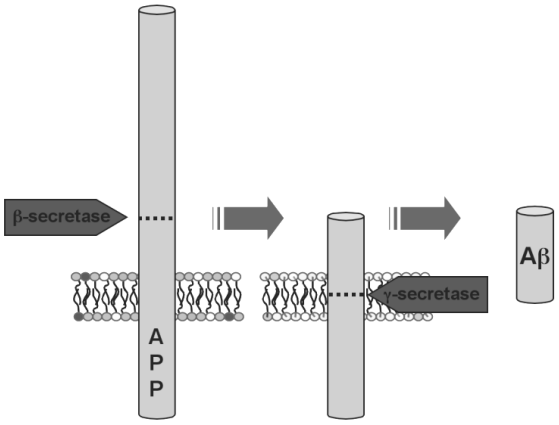


Figure 6.18 - *In vivo* generation of the Aβ peptide from Amyloid Precursor Protein (APP). Cleavage of APP by β-secretase forms the C-terminal membrane-bound APP fragment which is then cleaved by γ-secretase to yield the soluble Aβ peptide. γ-secretase cleavage is unspecific and can produce 39 to 43 amino acid long Aβ peptides.

protein mostly expressed in the synapses of neurons involved in synapse formation [148], neural plasticity [149] and copper metabolism [150]. Unspecific APP processing produces distinct Aβ peptides. The major forms are the peptides with 40 ($A\beta_{1-40}$) and 42 ($A\beta_{1-42}$) amino acid residues [143, 151-153]. $A\beta_{1-42}$ is significantly more amyloidogenic than $A\beta_{1-40}$, initiating amyloid formation at lower concentrations, with higher rates [154], resulting in different structures [155] and whose metabolism imbalancement is closely related to disease [156].

The most accepted hypothesis for the origin of the disease is the amyloid cascade hypothesis, which postulates that the disease is primarily the outcome of the toxicity and physiological imbalancement originating from amyloid formation and deposition: oxidative stress, calcium homeostasis imbalancement [157], apoptosis [158], mitochondrial enzyme dysfunction [159], inflammation [160], neurovascular dysfunction and cell cycle abnormalities. Even though insoluble Aβ amyloid plaques are the most characteristic pathological hallmark of the disease, accumulating evidence suggests that the disease-causing agent are not Aβ fibrils but rather oligomers formed during the amyloidogenesis pathway [107, 161]. Aβ oligomers trigger specific neuronal dysfunction [162-164], disrupt long-term potentiation [165] and impair memory [166]. The highest toxicity is

associated with Aβ oligomers formed by 2-50 monomers [157].The relevance of $A\beta_{1-42}$ has been consistently associated with its propensity of forming a diverse range of oligomeric structures which are stable, unlike those forming from $A\beta_{1-40}$ [167-171]. Nevertheless, both peptides eventually polymerize into structurally indistinguishable fibrils [172-173].

Aβ coordination chemistry is rich and it can bind different metal ions which influence its structure and amyloidogenic properties (Figure 6.19). Copper and zinc – two metal ions released during neuronal activity – bind Aβ plaques directly [128, 174]. Iron accumulates in the vicinity, in plaque-associated neuritic processes containing ferritin [175-176]. Overall, in the brain regions where amyloid deposition occurs Alzheimer's and Parkinson's disease (cortex in Alzheimer's and *substantia nigra* in Parkinson's, respectively) iron and zinc levels are found increased [177-178] while

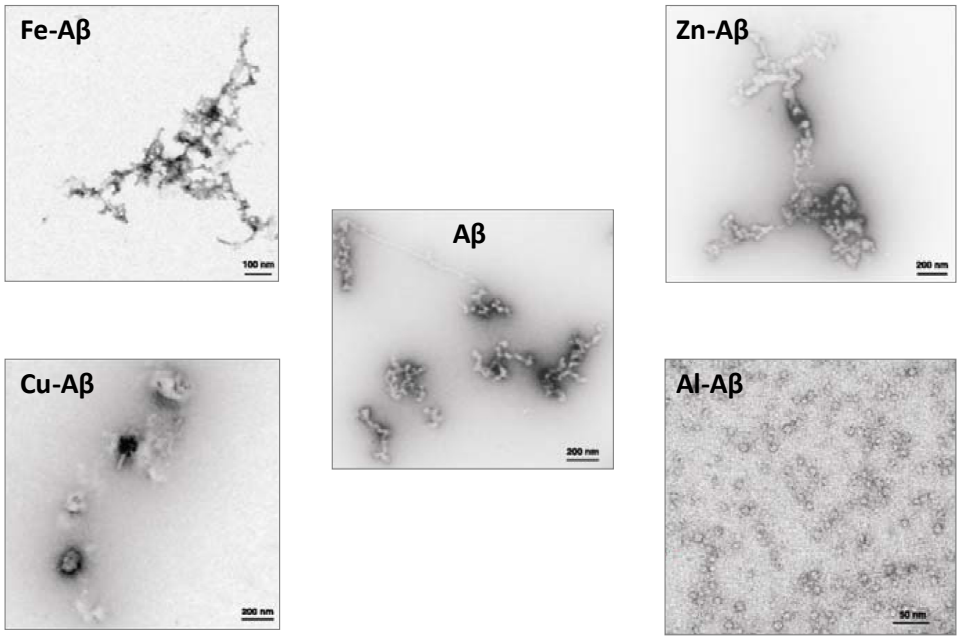


Figure 6.19 – Metal-dependent $A\beta_{1-42}$ conformational changes. Apo $A\beta_{1-42}$ is predominantly fibrillar. Incubation with iron, copper or zinc produces different types of amorphous oligomeric-like structures. Aluminum stabilizes small oligomeric forms, known to be the toxic species associated with Alzheimer's disease. From [23].

copper levels are decreased [179]. Copper-derived oxidative stress contributes significantly to disease onset [180]. Iron-related reactive oxygen species generation has been reported *in vitro* [181]. This may involve modifications of Aβ itself at methionine [182], histidine [183], glutamate [184] and tyrosine [185] residues. In the case of tyrosines, a frequent modification involves the formation of protease-resistant dityrosine cross-links, which favor further oligomerization [186]. Zinc binding inhibits redox reactions but induces Aβ aggregation [14].

The zinc-associated Aβ aggregates are especially relevant in the physiological context of Aβ deposition. Glutamatergic synapses – one of the sites of Aβ accumulation – release up to micromolar zinc transients (millisecond timescale) in response to upstream neurotransmission-associated stimuli, the so-called "zinc bursts". *In vitro*, these concentrations conditions induce fast Aβ aggregation in the form of prefibrillar oligomers – the putative toxic species in Alzheimer's disease – precluding proper fibril formation (Figure 6.20). Accordingly, it has been proposed that zinc homeostasis imbalancement contributes to disease onset by triggering an "amyloid cascade" of perpetuating self-seeding toxic oligomers [135]. Disease progression correlates with an increase in overall zinc levels which suggests enhanced metal-induced amyloid formation [178].

The chemical biology of glutamatergic synapse features the simultaneous availability of free zinc and copper, a unique coincidence at least in humans [179] which may explain the initial aggregation of Aβ at this site in Alzheimer's disease (Figure 6.21). The zinc in the cytoplasm of glutamatergic neurons is pumped by the ZnT3 zinc transporter to intracellular vesicles which also store glutamate. Upon stimulation, the secretory vesicles fuse with the pre-synaptic membrane and release its content into the synaptic space [187]. Zinc seems to prevent excess stimulation by glutamate by inhibiting the post-synaptic glutamate NMDA receptor (NMDAR). NMDAR

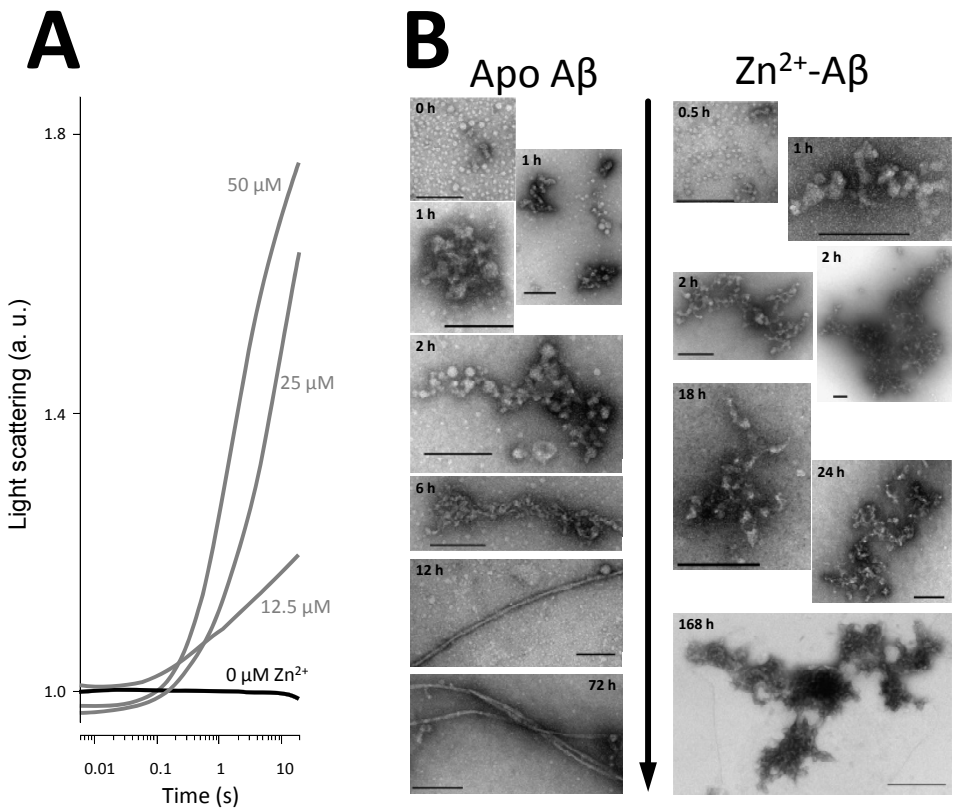


Figure 6.20 - Zn^{2+}-induced fast Aβ oligomerization. (A) Aβ aggregation kinetics as determined by stopped-flow light scattering measurements in the presence of different Zn^{2+} concentrations. Even sub maximal zinc bursts (which reach hundreds μM *in vivo*) promote fast Aβ aggregation. 25 μM Aβ. (B) Aβ aggregate morphology time course from transmission electron microscopy imaging. The fast Zn^{2+}-induced aggregation is associated to oligomeric structure assembly, dissimilarly from the fibrils forming with apo Aβ. From [135].

activation triggers post-synaptic copper secretion involving the Menkes ATPase [188]. In the synapse, copper can function as an electron acceptor promoting the reaction of nitric oxide with thiols, thus controlling NMDAR nitrosylation, a mechanism modulating NMDAR function [188]. Metallothionein-3 (MT3) released by astrocytes removes the metals from the synapse [189].

Metal ions like zinc copper and iron fulfill neurotransmission roles in neurobiological processes. In addition, neurodegenerative diseases like

Alzheimer's and Parkinson's are characterized by perturbations in metal homeostasis: iron levels are increased and copper and zinc are sequestered in amyloid deposits. The copper and zinc flux in the glutamatergic in the cortex and hippocampus is especially relevant because this is the site of long-term potentiation – the physiological basis of memory – and also the first amyloid deposition site in Alzheimer's disease. In neurotransmission events, free ionic or weakly protein bound (thus exchangeable) zinc and copper are released into the synaptic cleft. Labile intracellular iron modulates the expression of amyloid precursor protein [190]. In mouse models of ALS, knock-out of the metallochaperone donating copper to SOD1, CCS, has no effect on pathology. However, overexpression accelerates disease progression, pointing to a complex relationship of copper with the disease [191].

Metal mis-homeostasis has been suggested to trigger amyloid formation, as assessed by the significant reduction in amyloid plaques in mice deficient for zinc transporter ZnT3 [192]. Metal homeostatic mechanisms are affected by age. In animals, copper levels decrease from middle age onwards, affecting

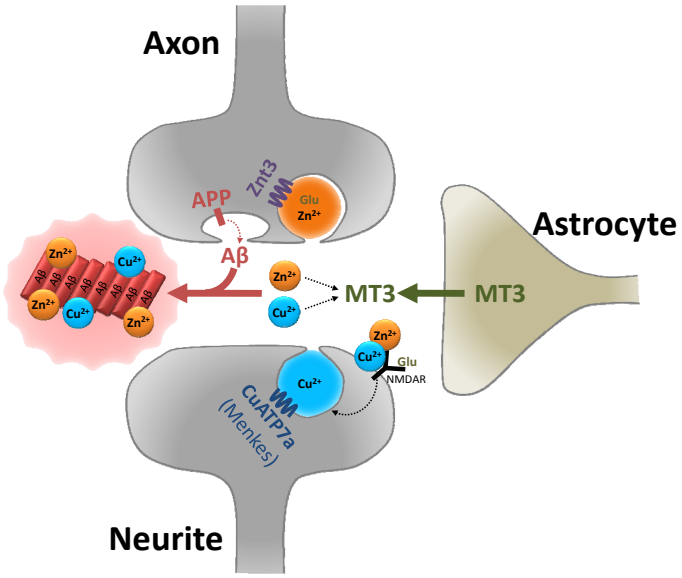


Figure 6.21 - Zinc, copper and Aβ in the glutamatergic synapse. From [179].

metal incorporation in copper-dependent proteins like cytochrome *c* oxidase, superoxide dismutase 1 – an important oxidative defense protein – and ceruloplasmin. Similarly, all animal species show increased iron levels in older ages, which affect proteins like ferritin and frataxin, whose misregulation is associated with neurodegeneration [179].

Although Aβ does not bind calcium, calcium-associated toxicity is one of the characteristics of Alzheimer's disease (Figure 6.22). Calcium homeostasis is essential for several signaling processes which are disrupted by Aβ. Aβ modulates the conductance of several membrane calcium channels [157, 193] including voltage-gated calcium channels, nicotinic acetylcholine channels and glutamate, dopamine, serotonin and intracellular inositol trisphosphate receptors through mechanisms mostly unknown. The calcium signaling perturbation induces excitotoxicity, neurotransmission perturbation, downstream signaling inhibition or aberrant stimulation. Aβ can also interact with membranes, partially disrupting their integrity and permeabilizing them to small molecules [194] or by forming integral membrane pores [195]. Sub-physiological calcium concentrations can favor

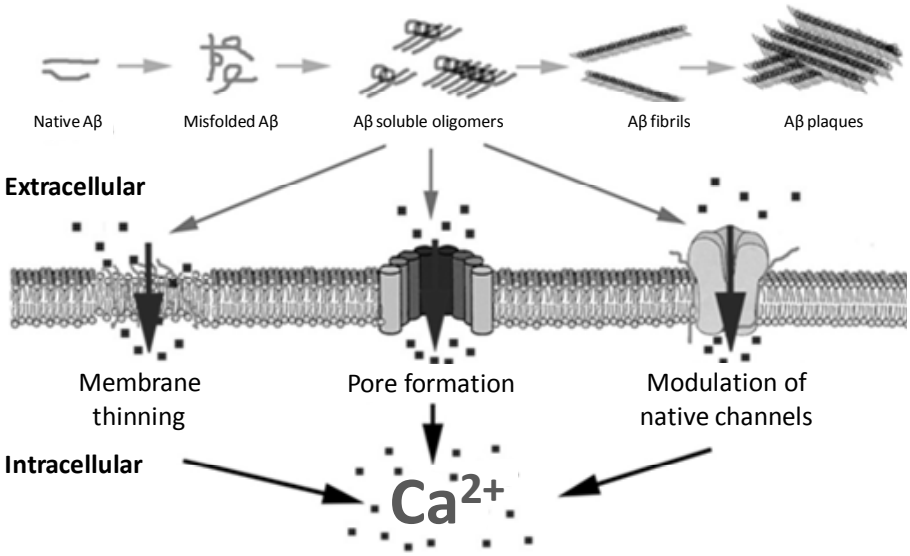


Figure 6.22 - Association of Aβ aggregation and calcium-associated cellular toxicity. From [157].

further Aβ generation [157] or bind Aβ and induce oligomerization and fibrillization processes [196-197].

6.4.2. Metal ions and Parkinson's disease

Parkinson's disease is characterized by the loss of dopaminergic neurons in the *substantia nigra* of the brain where intracellular amyloid deposits named Lewy bodies composed mainly of α-synuclein form in the surviving neurons. α-synuclein is a natively unstructured protein [198-200] involved in presynaptic vesicle homeostasis which can bind di- and trivalent metal ions. The affinity for Cu^{2+} and Fe^{3+} binding is relatively high. Protein function, including membrane association and interaction with cytoskeletal proteins is regulated by Ca^{2+} [201-202]. Most data on the roles of metal ions in α-synuclein amyloid formation come from *in vitro* experiments rather than from direct examination of living specimens. Although the mechanisms underlying α-synuclein fibrillization *in vivo* are not known, high temperature [203], low pH [203], polyamines [204] and metal ions can induce amyloid formation *in vitro*. Like in Alzheimer's, Parkinson's disease is characterized by altered metal homeostasis. The *substantia nigra* has increased iron and zinc levels and decreased copper levels [205-206]. Specifically in Lewy bodies, iron and aluminum are accumulated [207]. Iron accumulation perturbs its homeostatic mechanisms [208] and has been related to oxidative stress-related neurodegenerative damage in the disease [209], which is additionally relevant because of the decreased concentration of oxidative defense systems such as glutathione [206], as well as increased toxicity of α-synuclein amyloid oligomers [210].

Cu^{2+} and, to a less extent, Al^{3+} induce α-synuclein aggregation [211]. α-synuclein fibrillization is significantly accelerated in the presence of metal ions, namely Al^{3+}, Fe^{3+}, Cu^{2+}. Non-physiological metal ions like Co^{2+} and Cd^{2+} also have the same effect. To a small extent even Mn^{2+} also accelerates fibrillization and induces di-tyrosine cross-linking, pointing to a the capacity

of promoting oxidative modifications [139]. These effects should be based on the capacity of metal ions induce partial folding of natively unfolded α-synuclein [139]. However, Al^{3+} but not Zn^{2+} and Cu^{2+} induce a fibrillar structure distinct from the one of apo α-synuclein with higher β-sheet content and which does not favor iron Fenton reactivity [212]. The interest in assessing the effects of aluminum comes from the correlation of environmental exposure to aluminum and the accumulation of iron as well as aluminum in Lewy bodies [207]. The Al^{3+}-bound α-synuclein acquires a protease resistant conformation [213]. Upon fibrillization, aluminum is entrapped in a-synuclein amyloid fibrils [139]. Magnesium has a dual effect in α-synuclein fibrillization. Millimolar concentrations promptly trigger amyloid formation [214]. However, lower concentrations inhibit fibril formation, suggesting a conformational stabilizer role [215].

Despite the similar amyloid favoring properties, metal ions bind at distinct sites of α-synuclein. The C-terminus is the most common metal binding region because of the local residual structure. The site involving Asp119, Asp121, Glu126 and possibly phospho-Tyr125 binds Fe^{2+}, Mn^{2+}, Co^{2+} and Ni^{2+} with submillimolar affinity [216]. Copper binds at the N-terminus (K_d ~0.1 μM) [217] involving two binding site [218]. Methionine oxidation inhibits α-synuclein fibrillization, an effect which can be reverted by zinc binding [140].

Like Aβ, α-synuclein fibrillization pathways include a wide variety of oligomer structures [219-220], some of the molten globule type [203]. Metal ions induce distinct oligomeric structures [221]: Cu^{2+}, Fe^{3+} and Ni^{2+} yield apo-like 0.8-4 nm spherical oligomers; Mg^{2+}, Cd^{2+} and Zn^{2+} yield larger 5-8 nm spherical structures; Co^{2+} yields 22-30 nm wide annular oligomers; Ca^{2+} yields larger 70-90 nm wide annular oligomers. All oligomer types are toxic [220]. The toxicity of annular oligomers is derived, at least in part, to pore formation in membranes. Calcium binding occurs in the C-terminus of α-

synuclein with a K_d around 2-300 μM [202]. Ca^{2+} binding affects α-synuclein interaction with cytoskeletal proteins and favors oligomerization. α-synuclein deposition affects calcium homeostasis, activating voltage-gated calcium channels [222] and calcium-dependent proteases and determining mitochondrial dysfunction [223]. Like in Alzheimer's disease, the altered calcium homeostasis sets a feed-forward loop, unbalancing α-synuclein metabolism and favoring more aggregation [197].

6.4.3. Metal ions and prion diseases

Prions (proteinaceous infectious particles) are protein infectious agents responsible for transmissible spongiform encephelopathies (TSEs), which are amyloid neurodegenerative diseases affecting sheep, goats (Scrapie), cows (bovine spongiform encephalopathy), humans (Creutzfeldt-Jakob disease, Gerstmann-Sträussler-Scheinker syndrome, fatal familial insomnia) and other animals. The cellular prion protein (PrP^C) is a normal constituent of the brain, being most expressed at synapses [224], where it binds membranes through a glycophosphatidylinositol anchor. The native PrP^C conformation is monomeric, mostly unstructured and protease sensitive. It can spontaneously convert to an amyloidogenic – or scrapie [225] – conformation (PrP^{SC}) which is β-sheet rich, amyloid-prone, has low solubility and is highly protease resistant which initiates the autocatalytic amyloid formation process [226].

The prion protein has ~210 amino acid residues, its accurate size depending on the species. The C-terminus contains two short β-strands and three α-helices but the N-terminus is unstructured [227] and includes several PHGGSWGN octarepeats which bind copper in a pH dependent stoichiometry: two Cu^{2+} ions at pH 6 and four at pH 7.4. The binding at lower pH has a K_d in the low micromolar range. At higher pH the binding cooperativity is high and binding of the last two ions has a K_d < 100 nM [228]. The dissociation constants approach the cerebrospinal fluid copper

concentrations [229] (0.3 to 500 μM) [230] and binding is observed *in vivo* [231]. It induces PrP^{C} endocytosis. The octarepeat region can also accommodate Zn^{2+} in a competitive way. According to the Irving-Williams series, copper is the preferred ligand. It can be displaced by Zn^{2+} at nanomolar Cu^{2+} concentrations and high millimolar Zn^{2+} concentrations [232]. Other metal ligands are Fe^{2+}, Ni^{2+} and Mn^{2+} [233-234]. Upon conversion to the PrP^{SC} state, the octarepeat undergoes a conformational change where it becomes more compact [235].

Iron, copper and manganese levels are affected in the brains infected by PrP^{SC} and this is associated with prion protein conformational changes or increased oxidative stress generation [236-237], which may also perturb protein folding (Figure 6.23). Brain copper levels are lower in PrP^{SC} infected subjects [238]. Copper binding promptly and reversibly promotes a conformational change wherein prion protein becomes detergent insoluble and protease resistant without acquiring the scrapie amyloid conformation [239] and inhibiting fibrillization [240] and disease progression [241]. Copper chelation delays onset of disease [242]. Copper-bound PrP^{C} has superoxide dismutase (SOD)-like activity [243-244]. Infection with PrP^{SC} – with the concomitant conversion of PrP^{C} – is associated with a decrease in superoxide dismutase activity and increased oxidative stress. On the other hand, once PrP^{SC} is formed, copper greatly stabilizes it against proteinase K digestion [245]. Iron induces the formation of the protease-resistant state [246]. Zn^{2+} critical for PrP 106-126 aggregation and neurotoxicity [142]. Manganese is one of the most important modulators of PrP^{C} aggregation because it greatly enhances aggregation [238, 247] into the protease-resistant state [246, 248]. Like in other amyloid diseases, cytotoxicity is mostly associated with oligomeric structures [107].

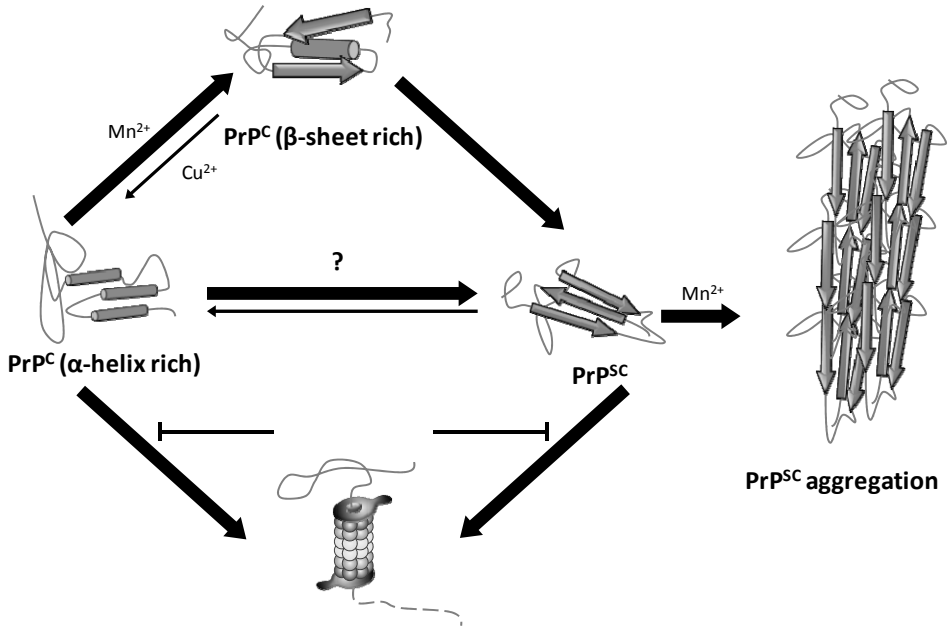


Figure 6.23 – Metal-induced conformational changes in the prion protein. PrP^{C} and PrP^{SC} may exist in equilibrium with each other. Most of the time PrP^{SC} is degraded and remains undetectable. PrP^{SC} levels can increase due to proteasome activity inhibition by metal ions like Mn^{2+}, Cd^{2+} or Zn^{2+}, by binding of amyloid-favoring Mn^{2+} to PrP^{SC} or by conformational fluctuations of PrP^{C} to β-sheet enriched conformations which can convert to PrP^{SC}. Cu^{2+} binding stabilizes the non amyloidogenic state. Once a PrP^{SC} nucleus is formed, recruitment of more prion protein to the amyloid deposits proceeds through conformational templating. From [233].

6.5. References

1. Andreini C, Bertini I, Cavallaro G, Holliday GL & Thornton JM (2008) Metal ions in biological catalysis: from enzyme databases to general principles. *J Biol Inorg Chem* **13**, 1205-1218
2. Dupont CL, Yang S, Palenik B & Bourne PE (2006) Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. *Proc Natl Acad Sci U S A* **103**, 17822-17827
3. Fraústo da Silva JJR & Williams JRP (2001) *The Biological Chemistry of the Elements*. Oxford University Press.
4. Saito MA, Sigman DM & Morel FMM (2003) The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean–Proterozoic boundary? *Inorg Chim Acta* **356**, 308-318
5. Gray HB (2003) Biological inorganic chemistry at the beginning of the 21st century. *Proc Natl Acad Sci U S A* **100**, 3563-3568
6. Wittung-Stafshede P (2002) Role of cofactors in protein folding. *Acc Chem Res* **35**, 201-208
7. Waldron KJ, Rutherford JC, Ford D & Robinson NJ (2009) Metalloproteins and metal sensing. *Nature* **460**, 823-830
8. Wachtershauser G (1988) Before enzymes and templates: theory of surface metabolism. *Microbiol Rev* **52**, 452-484
9. Koch AL & Schmidt TM (1991) The first cellular bioenergetic process: primitive generation of a proton-motive force. *J Mol Evol* **33**, 297-304
10. Permyakov E (2009) *Metalloproteomics*. 1st edition edn. John Wiley & Sons, Hoboken, NJ.

11. Papoian GA, Ulander J & Wolynes PG (2003) Role of water mediated interactions in protein-protein recognition landscapes. *J Am Chem Soc* **125**, 9170-9178
12. Wang J & Verkhivker GM (2003) Energy landscape theory, funnels, specificity, and optimal criterion of biomolecular binding. *Phys Rev Lett* **90**, 188101
13. Rae TD, Schmidt PJ, Pufahl RA, Culotta VC & O'Halloran TV (1999) Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase. *Science* **284**, 805-808
14. Frederickson CJ, Koh JY & Bush AI (2005) The neurobiology of zinc in health and disease. *Nature reviews* **6**, 449-462
15. Knight MR, Campbell AK, Smith SM & Trewavas AJ (1991) Recombinant aequorin as a probe for cytosolic free Ca2+ in Escherichia coli. *FEBS Lett* **282**, 405-408
16. Jones HE, Holland IB, Baker HL & Campbell AK (1999) Slow changes in cytosolic free Ca2+ in Escherichia coli highlight two putative influx mechanisms in response to changes in extracellular calcium. *Cell Calcium* **25**, 265-274
17. Yamasaki S, Sakata-Sogawa K, Hasegawa A, Suzuki T, Kabu K, Sato E, Kurosaki T, Yamashita S, Tokunaga M, Nishida K, et al. (2007) Zinc is a novel intracellular second messenger. *J Cell Biol* **177**, 637-645
18. Carafoli E (2002) Calcium signaling: a tale for all seasons. *Proc Natl Acad Sci U S A* **99**, 1115-1122
19. Rosenzweig AC (2002) Metallochaperones: bind and deliver. *Chem Biol* **9**, 673-677
20. Lill R & Kispal G (2000) Maturation of cellular Fe-S proteins: an essential function of mitochondria. *Trends Biochem Sci* **25**, 352-356
21. Krebs C, Agar JN, Smith AD, Frazzon J, Dean DR, Huynh BH & Johnson MK (2001) IscA, an alternate scaffold for Fe-S cluster biosynthesis. *Biochemistry* **40**, 14069-14080
22. Sigel A & Sigel H (2009) Metallothioneins and Related Chelators. In *Metal Ions in Life Sciences* (Sigel RKO, ed. RSC Publishing, Cambridge.
23. Zatta P, Drago D, Bolognin S & Sensi SL (2009) Alzheimer's disease, metal ions and metal homeostatic therapy. *Trends Pharmacol Sci* **30**, 346-355
24. Tottey S, Waldron KJ, Firbank SJ, Reale B, Bessant C, Sato K, Cheek TR, Gray J, Banfield MJ, Dennison C, et al. (2008) Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. *Nature* **455**, 1138-1142
25. Irving HMNH & Williams RJP (1953) The stability of transition-metal complexes. *J Chem Soc*, 3192-3210
26. Outten CE & O'Halloran TV (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-2492
27. Robinson CR, Liu Y, Thomson JA, Sturtevant JM & Sligar SG (1997) Energetics of heme binding to native and denatured states of cytochrome b562. *Biochemistry* **36**, 16141-16146
28. Wittung-Stafshede P, Lee JC, Winkler JR & Gray HB (1999) Cytochrome b562 folding triggered by electron transfer: approaching the speed limit for formation of a four-helix-bundle protein. *Proc Natl Acad Sci U S A* **96**, 6587-6590
29. Wittung-Stafshede P, Hill MG, Gomez E, Bilio AJD, Karlsson BG, Leckner J, Winkler JR, Gray HB & Malmström BG (1998) Reduction potentials of blue and purple copper proteins in their unfolded states: a closer look at rack-induced coordination. *J Biol Inorg Chem* **3**, 367-370
30. Pozdnyakova I, Guidry J & Wittung-Stafshede P (2001) Probing copper ligands in denatured Pseudomonas aeruginosa azurin: unfolding His117Gly and His46Gly mutants. *J Biol Inorg Chem* **6**, 182-188
31. Wittung-Stafshede P, Malmstrom BG, Sanders D, Fee JA, Winkler JR & Gray HB (1998) Effect of redox state on the folding free energy of a thermostable electron-transfer metalloprotein: the CuA domain of cytochrome oxidase from Thermus thermophilus. *Biochemistry* **37**, 3172-3177
32. Gomes CM & Wittung-Stafshede P (2010) Metal Ions, Protein Folding ans Conformational States - An introduction. In *Protein Folding and Metal Ions: Mechanisms, Biology and Disease* (Gomes CM & Wittung-Stafshede P, eds). CRC Press, USA.
33. Varhac R, Antalik M & Bano M (2004) Effect of temperature and guanidine hydrochloride on ferrocytochrome c at neutral pH. *J Biol Inorg Chem* **9**, 12-22
34. Tezcan FA, Findley WM, Crane BR, Ross SA, Lyubovitsky JG, Gray HB & Winkler JR (2002) Using deeply trapped intermediates to map the cytochrome c folding landscape. *Proc Natl Acad Sci U S A* **99**, 8626-8630

35. Lyubovitsky JG, Gray HB & Winkler JR (2002) Mapping the cytochrome C folding landscape. *J Am Chem Soc* **124**, 5481-5485
36. Fisher WR, Taniuchi H & Anfinsen CB (1973) On the role of heme in the formation of the structure of cytochrome c. *J Biol Chem* **248**, 3188-3195
37. Yamanaka M, Mita H, Yamamoto Y & Sambongi Y (2009) Heme is not required for Aquifex aeolicus cytochrome c(555) polypeptide folding. *Biosci Biotechnol Biochem* **73**, 2022-2025
38. Borgia A, Gianni S, Brunori M & Travaglini-Allocatelli C (2008) Fast folding kinetics and stabilization of apo-cytochrome c. *FEBS Lett* **582**, 1003-1007
39. Aisen P & Listowsky I (1980) Iron transport and storage proteins. *Annu Rev Biochem* **49**, 357-393
40. Lin LN, Mason AB, Woodworth RC & Brandts JF (1994) Calorimetric studies of serum transferrin and ovotransferrin. Estimates of domain interactions, and study of the kinetic complexities of ferric ion binding. *Biochemistry* **33**, 1881-1888
41. Lin LN, Mason AB, Woodworth RC & Brandts JF (1993) Calorimetric studies of the N-terminal half-molecule of transferrin and mutant forms modified near the Fe(3+)-binding site. *The Biochemical journal* **293 (Pt 2)**, 517-522
42. Andreini C, Banci L, Bertini I & Rosato A (2006) Zinc through the three domains of life. *J Proteome Res* **5**, 3173-3178
43. Andreini C, Banci L, Bertini I & Rosato A (2006) Counting the zinc-proteins encoded in the human genome. *J Proteome Res* **5**, 196-201
44. Lander ES & Linton LM & Birren B & Nusbaum C & Zody MC & Baldwin J & Devon K & Dewar K & Doyle M & FitzHugh W, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921
45. Berg JM & Godwin HA (1997) Lessons from zinc-binding peptides. *Annu Rev Biophys Biomol Struct* **26**, 357-371
46. Frankel AD, Berg JM & Pabo CO (1987) Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc Natl Acad Sci U S A* **84**, 4841-4845
47. Michael SF, Kilfoil VJ, Schmidt MH, Amann BT & Berg JM (1992) Metal binding and folding properties of a minimalist Cys2His2 zinc finger peptide. *Proc Natl Acad Sci U S A* **89**, 4796-4800
48. Laity JH, Lee BM & Wright PE (2001) Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol* **11**, 39-46
49. Potter SZ, Zhu H, Shaw BF, Rodriguez JA, Doucette PA, Sohn SH, Durazo A, Faull KF, Gralla EB, Nersissian AM, et al. (2007) Binding of a single zinc ion to one subunit of copper-zinc superoxide dismutase apoprotein substantially influences the structure and stability of the entire homodimeric protein. *J Am Chem Soc* **129**, 4575-4583
50. Yang Y & Zhou HM (2001) Effect of zinc ions on conformational stability of yeast alcohol dehydrogenase. *Biochemistry (Mosc)* **66**, 47-54
51. Le WP, Yan SX, Zhang YX & Zhou HM (1996) Acid-induced folding of yeast alcohol dehydrogenase under low pH conditions. *J Biochem* **119**, 674-679
52. Le WP, Yan SX, Li S, Zhong HN & Zhou HM (1996) Alkaline unfolding and salt-induced folding of yeast alcohol dehydrogenase under high pH conditions. *Int J Pept Protein Res* **47**, 484-490
53. Nar H, Messerschmidt A, Huber R, van de Kamp M & Canters GW (1992) Crystal structure of Pseudomonas aeruginosa apo-azurin at 1.85 A resolution. *FEBS Lett* **306**, 119-124
54. Nar H, Messerschmidt A, Huber R, van de Kamp M & Canters GW (1991) Crystal structure analysis of oxidized Pseudomonas aeruginosa azurin at pH 5.5 and pH 9.0. A pH-induced conformational transition involves a peptide bond flip. *J Mol Biol* **221**, 765-772
55. Wittung-Stafshede P (2004) Role of cofactors in folding of the blue-copper protein azurin. *Inorg Chem* **43**, 7926-7933
56. De Filippis V, Vassiliev VB, Beltramini M, Fontana A, Salvato B & Gaitskhoki VS (1996) Evidence for the molten globule state of human apo-ceruloplasmin. *Biochim Biophys Acta* **1297**, 119-123
57. Sedlak E, Zoldak G & Wittung-Stafshede P (2008) Role of copper in thermal stability of human ceruloplasmin. *Biophys J* **94**, 1384-1391
58. Bento I, Peixoto C, Zaitsev VN & Lindley PF (2007) Ceruloplasmin revisited: structural and functional roles of various metal cation-binding sites. *Acta Crystallogr D Biol Crystallogr* **63**, 240-248
59. Busse R & Mulsch A (1990) Calcium-dependent nitric oxide synthesis in endothelial cytosol is mediated by calmodulin. *FEBS Lett* **265**, 133-136

60. Lee YC & Wolff J (1982) Two opposing effects of calmodulin on microtubule assembly depend on the presence of microtubule-associated proteins. *J Biol Chem* **257**, 6306-6310
61. Uboha NV, Flajolet M, Nairn AC & Picciotto MR (2007) A calcium- and calmodulin-dependent kinase Ialpha/microtubule affinity regulating kinase 2 signaling cascade mediates calcium-dependent neurite outgrowth. *J Neurosci* **27**, 4413-4423
62. Kretsinger RH & Nockolds CE (1973) Carp muscle calcium-binding protein. II. Structure determination and general description. *J Biol Chem* **248**, 3313-3326
63. Laberge M, Wright WW, Sudhakar K, Liebman PA & Vanderkooi JM (1997) Conformational effects of calcium release from parvalbumin: comparison of computational simulations with spectroscopic investigations. *Biochemistry* **36**, 5363-5371
64. Permyakov EA, Ostrovsky AV, Burstein EA, Pleshanov PG & Gerday C (1985) Parvalbumin conformers revealed by steady-state and time-resolved fluorescence spectroscopy. *Arch Biochem Biophys* **240**, 781-791
65. Permyakov SE, Bakunts AG, Denesyuk AI, Knyazeva EL, Uversky VN & Permyakov EA (2008) Apo-parvalbumin as an intrinsically disordered protein. *Proteins* **72**, 822-836
66. Henzl MT & Tanner JJ (2007) Solution structure of Ca2+-free rat beta-parvalbumin (oncomodulin). *Protein Sci* **16**, 1914-1926
67. Cox JA, Winge DR & Stein EA (1979) Calcium, magnesium and the conformation of parvalbumin during muscular activity. *Biochimie* **61**, 601-605
68. Forge V, Wijesinha RT, Balbach J, Brew K, Robinson CV, Redfield C & Dobson CM (1999) Rapid collapse and slow structural reorganisation during the refolding of bovine alpha-lactalbumin. *J Mol Biol* **288**, 673-688
69. Ikeguchi M, Kuwajima K & Sugai S (1986) Ca2+-induced alteration in the unfolding behavior of alpha-lactalbumin. *J Biochem* **99**, 1191-1201
70. Bushmarina NA, Blanchet CE, Vernier G & Forge V (2006) Cofactor effects on the protein folding reaction: acceleration of alpha-lactalbumin refolding by metal ions. *Protein Sci* **15**, 659-671
71. Rao KR & Brew K (1989) Calcium regulates folding and disulfide-bond formation in alpha-lactalbumin. *Biochem Biophys Res Commun* **163**, 1390-1396
72. Nitta K, Tsuge H, Shimazaki K & Sugai S (1988) Calcium-binding lysozymes. *Biol Chem Hoppe Seyler* **369**, 671-675
73. Nitta K, Tsuge H, Sugai S & Shimazaki K (1987) The calcium-binding property of equine lysozyme. *FEBS Lett* **223**, 405-408
74. Permyakov SE, Khokhlova TI, Nazipova AA, Zhadan AP, Morozova-Roche LA & Permyakov EA (2006) Calcium-binding and temperature induced transitions in equine lysozyme: new insights from the pCa-temperature "phase diagrams". *Proteins* **65**, 984-998
75. Astbury WT & Dickinson S (1935) The X-ray interpretation of denaturation and the structure of the seed globulins. *The Biochemical journal* **29**, 2351-2360 2351
76. Eanes ED & Glenner GG (1968) X-ray diffraction studies on amyloid filaments. *J Histochem Cytochem* **16**, 673-677
77. Sunde M & Blake CC (1998) From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Q Rev Biophys* **31**, 1-39
78. Tycko R (2004) Progress towards a molecular-level structural understanding of amyloid fibrils. *Curr Opin Struct Biol* **14**, 96-103
79. Ross CA & Poirier MA (2004) Protein aggregation and neurodegenerative disease. *Nat Med* **10 Suppl**, S10-17
80. Clark PL (2004) Protein folding in the cell: reshaping the folding funnel. *Trends Biochem Sci* **29**, 527-534
81. Hartl FU & Hayer-Hartl M (2009) Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol* **16**, 574-581
82. Guijarro JI, Sunde M, Jones JA, Campbell ID & Dobson CM (1998) Amyloid fibril formation by an SH3 domain. *Proc Natl Acad Sci U S A* **95**, 4224-4228
83. Gazit E (2002) The "Correctly Folded" state of proteins: is it a metastable state? *Angew Chem Int Ed Engl* **41**, 257-259
84. Meersman F & Dobson CM (2006) Probing the pressure-temperature stability of amyloid fibrils provides new insights into their molecular properties. *Biochim Biophys Acta* **1764**, 452-460

85. Wisniewski T, Ghiso J & Frangione B (1994) Alzheimer's disease and soluble A beta. *Neurobiol Aging* **15**, 143-152
86. Li L & Lindquist S (2000) Creating a protein-based element of inheritance. *Science* **287**, 661-664
87. Maury CP (2009) The emerging concept of functional amyloid. *J Intern Med* **265**, 329-334
88. Dobson CM (1999) Protein misfolding, evolution and disease. *Trends Biochem Sci* **24**, 329-332
89. Chiti F & Dobson CM (2009) Amyloid formation by globular proteins under native conditions. *Nat Chem Biol* **5**, 15-22
90. Hu X, Crick SL, Bu G, Frieden C, Pappu RV & Lee JM (2009) Amyloid seeds formed by cellular uptake, concentration, and aggregation of the amyloid-beta peptide. *Proc Natl Acad Sci U S A* **106**, 20324-20329
91. Jackson GS, Hosszu LL, Power A, Hill AF, Kenney J, Saibil H, Craven CJ, Waltho JP, Clarke AR & Collinge J (1999) Reversible conversion of monomeric human prion protein between native and fibrilogenic conformations. *Science* **283**, 1935-1937
92. Lashuel HA, Lai Z & Kelly JW (1998) Characterization of the transthyretin acid denaturation pathways by analytical ultracentrifugation: implications for wild-type, V30M, and L55P amyloid fibril formation. *Biochemistry* **37**, 17851-17864
93. Lee HJ, Khoshaghideh F, Patel S & Lee SJ (2004) Clearance of alpha-synuclein oligomeric intermediates via the lysosomal degradation pathway. *J Neurosci* **24**, 1888-1896
94. Armen RS, DeMarco ML, Alonso DO & Daggett V (2004) Pauling and Corey's alpha-pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease. *Proc Natl Acad Sci U S A* **101**, 11622-11627
95. Cannon MJ, Williams AD, Wetzel R & Myszka DG (2004) Kinetic analysis of beta-amyloid fibril elongation. *Anal Biochem* **328**, 67-75
96. Esler WP, Stimson ER, Jennings JM, Vinters HV, Ghilardi JR, Lee JP, Mantyh PW & Maggio JE (2000) Alzheimer's disease amyloid propagation by a template-dependent dock-lock mechanism. *Biochemistry* **39**, 6288-6295
97. Uversky VN (2010) Mysterious oligomerization of the amyloidogenic proteins. *FEBS J* **277**, 2940-2953
98. Morris AM, Watzky MA & Finke RG (2009) Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *Biochim Biophys Acta* **1794**, 375-397
99. Zhang J & Muthukumar M (2009) Simulations of nucleation and elongation of amyloid fibrils. *J Chem Phys* **130**, 035102
100. Baskakov IV (2007) Branched chain mechanism of polymerization and ultrastructure of prion protein amyloid fibrils. *FEBS J* **274**, 3756-3765
101. Wegner A & Savko P (1982) Fragmentation of actin filaments. *Biochemistry* **21**, 1909-1913
102. Makin OS & Serpell LC (2005) X-ray diffraction studies of amyloid structure. *Methods Mol Biol* **299**, 67-80
103. Sunde M & Blake C (1997) The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv Protein Chem* **50**, 123-159
104. Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB & Blake CC (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J Mol Biol* **273**, 729-739
105. Kayed R & Glabe CG (2006) Conformation-dependent anti-amyloid oligomer antibodies. *Methods Enzymol* **413**, 326-344
106. Kayed R, Head E, Sarsoza F, Saing T, Cotman CW, Necula M, Margol L, Wu J, Breydo L, Thompson JL, et al. (2007) Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol Neurodegener* **2**, 18
107. Kayed R, Head E, Thompson JL, McIntire TM, Milton SC, Cotman CW & Glabe CG (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**, 486-489
108. David DC, Ollikainen N, Trinidad JC, Cary MP, Burlingame AL & Kenyon C (2010) Widespread Protein Aggregation as an Inherent Part of Aging in C. elegans. *PLoS Biol* **8**
109. Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* **75**, 333-366
110. Merlini G & Bellotti V (2003) Molecular mechanisms of amyloidosis. *N Engl J Med* **349**, 583-596

111. Uversky VN (2007) Neuropathology, biochemistry, and biophysics of alpha-synuclein aggregation. *J Neurochem* **103**, 17-37

112. Saraiva MJ (2001) Transthyretin amyloidosis: a tale of weak interactions. *FEBS Lett* **498**, 201-203

113. Singleton A, Gwinn-Hardy K, Sharabi Y, Li ST, Holmes C, Dendi R, Hardy J, Crawley A & Goldstein DS (2004) Association between cardiac denervation and parkinsonism caused by alpha-synuclein gene triplication. *Brain* **127**, 768-772

114. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M, Peuralinna T, Dutra A, Nussbaum R, et al. (2003) alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841

115. Verdone G, Corazza A, Viglino P, Pettirossi F, Giorgetti S, Mangione P, Andreola A, Stoppini M, Bellotti V & Esposito G (2002) The solution structure of human beta2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci* **11**, 487-499

116. Shikama Y, Kitazawa J, Yagihashi N, Uehara O, Murata Y, Yajima N, Wada R & Yagihashi S (2010) Localized amyloidosis at the site of repeated insulin injection in a diabetic patient. *Intern Med* **49**, 397-401

117. Tan SY & Pepys MB (1994) Amyloidosis. *Histopathology* **25**, 403-414

118. Caughey B & Lansbury PT (2003) Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu Rev Neurosci* **26**, 267-298

119. Lashuel HA, Hartley D, Petre BM, Walz T & Lansbury PT, Jr. (2002) Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature* **418**, 291

120. Ferreira ST, Vieira MN & De Felice FG (2007) Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases. *IUBMB Life* **59**, 332-345

121. Nimmrich V, Grimm C, Draguhn A, Barghorn S, Lehmann A, Schoemaker H, Hillen H, Gross G, Ebert U & Bruehl C (2008) Amyloid beta oligomers (A beta(1-42) globulomer) suppress spontaneous synaptic activity by inhibition of P/Q-type calcium currents. *J Neurosci* **28**, 788-797

122. Nixon RA (2006) Autophagy in neurodegenerative disease: friend, foe or turncoat? *Trends Neurosci* **29**, 528-535

123. Powers ET, Morimoto RI, Dillin A, Kelly JW & Balch WE (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Annu Rev Biochem* **78**, 959-991

124. Barnham KJ, Ciccotosto GD, Tickler AK, Ali FE, Smith DG, Williamson NA, Lam YH, Carrington D, Tew D, Kocak G, et al. (2003) Neurotoxic, redox-competent Alzheimer's beta-amyloid is released from lipid membrane by methionine oxidation. *J Biol Chem* **278**, 42959-42965

125. Cuajungco MP, Goldstein LE, Nunomura A, Smith MA, Lim JT, Atwood CS, Huang X, Farrag YW, Perry G & Bush AI (2000) Evidence that the beta-amyloid plaques of Alzheimer's disease represent the redox-silencing and entombment of abeta by zinc. *J Biol Chem* **275**, 19439-19442

126. Huang X, Atwood CS, Hartshorn MA, Multhaup G, Goldstein LE, Scarpa RC, Cuajungco MP, Gray DN, Lim J, Moir RD, et al. (1999) The A beta peptide of Alzheimer's disease directly produces hydrogen peroxide through metal ion reduction. *Biochemistry* **38**, 7609-7616

127. Huang X, Cuajungco MP, Atwood CS, Hartshorn MA, Tyndall JD, Hanson GR, Stokes KC, Leopold M, Multhaup G, Goldstein LE, et al. (1999) Cu(II) potentiation of alzheimer abeta neurotoxicity. Correlation with cell-free hydrogen peroxide production and metal reduction. *J Biol Chem* **274**, 37111-37116

128. Opazo C, Huang X, Cherny RA, Moir RD, Roher AE, White AR, Cappai R, Masters CL, Tanzi RE, Inestrosa NC, et al. (2002) Metalloenzyme-like activity of Alzheimer's disease beta-amyloid. Cu-dependent catalytic conversion of dopamine, cholesterol, and biological reducing agents to neurotoxic H(2)O(2). *J Biol Chem* **277**, 40302-40308

129. Macario AJ & Conway de Macario E (2005) Sick chaperones, cellular stress, and disease. *N Engl J Med* **353**, 1489-1501

130. Muchowski PJ & Wacker JL (2005) Modulation of neurodegeneration by molecular chaperones. *Nature reviews* **6**, 11-22

131. Almeida CG, Takahashi RH & Gouras GK (2006) Beta-amyloid accumulation impairs multivesicular body sorting by inhibiting the ubiquitin-proteasome system. *J Neurosci* **26**, 4277-4288

132. Calabrese MF, Eakin CM, Wang JM & Miranker AD (2008) A regulatable switch mediates self-association in an immunoglobulin fold. *Nat Struct Mol Biol* **15**, 965-971

133. Eakin CM, Knight JD, Morgan CJ, Gelfand MA & Miranker AD (2002) Formation of a copper specific binding site in non-native states of beta-2-microglobulin. *Biochemistry* **41**, 10646-10656

134. Huang X, Atwood CS, Moir RD, Hartshorn MA, Tanzi RE & Bush AI (2004) Trace metal contamination initiates the apparent auto-aggregation, amyloidosis, and oligomerization of Alzheimer's Abeta peptides. *J Biol Inorg Chem* **9**, 954-960
135. Noy D, Solomonov I, Sinkevich O, Arad T, Kjaer K & Sagi I (2008) Zinc-amyloid beta interactions on a millisecond time-scale stabilize non-fibrillar Alzheimer-related species. *J Am Chem Soc* **130**, 1376-1383
136. Talmard C, Bouzan A & Faller P (2007) Zinc binding to amyloid-beta: isothermal titration calorimetry and Zn competition experiments with Zn sensors. *Biochemistry* **46**, 13658-13666
137. Talmard C, Guilloreau L, Coppel Y, Mazarguil H & Faller P (2007) Amyloid-beta peptide forms monomeric complexes with Cu(II) and Zn(II) prior to aggregation. *Chembiochem* **8**, 163-165
138. Talmard C, Leuma Yona R & Faller P (2009) Mechanism of zinc(II)-promoted amyloid formation: zinc(II) binding facilitates the transition from the partially alpha-helical conformer to aggregates of amyloid beta protein(1-28). *J Biol Inorg Chem* **14**, 449-455
139. Uversky VN, Li J & Fink AL (2001) Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. A possible molecular NK between Parkinson's disease and heavy metal exposure. *J Biol Chem* **276**, 44284-44296
140. Yamin G, Glaser CB, Uversky VN & Fink AL (2003) Certain metals trigger fibrillation of methionine-oxidized alpha-synuclein. *J Biol Chem* **278**, 27630-27635
141. Yamin G, Munishkina LA, Karymov MA, Lyubchenko YL, Uversky VN & Fink AL (2005) Forcing nonamyloidogenic beta-synuclein to fibrillate. *Biochemistry* **44**, 9096-9107
142. Jobling MF, Huang X, Stewart LR, Barnham KJ, Curtain C, Volitakis I, Perugini M, White AR, Cherny RA, Masters CL, et al. (2001) Copper and zinc binding modulates the aggregation and neurotoxic properties of the prion peptide PrP106-126. *Biochemistry* **40**, 8073-8084
143. Masters CL, Simms G, Weinman NA, Multhaup G, McDonald BL & Beyreuther K (1985) Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc Natl Acad Sci U S A* **82**, 4245-4249
144. Dickson DW (1997) The pathogenesis of senile plaques. *J Neuropathol Exp Neurol* **56**, 321-339
145. Selkoe DJ (1999) Translating cell biology into therapeutic advances in Alzheimer's disease. *Nature* **399**, A23-31
146. Mudher A & Lovestone S (2002) Alzheimer's disease-do tauists and baptists finally shake hands? *Trends Neurosci* **25**, 22-26
147. Blennow K, de Leon MJ & Zetterberg H (2006) Alzheimer's disease. *Lancet* **368**, 387-403
148. Priller C, Bauer T, Mitteregger G, Krebs B, Kretzschmar HA & Herms J (2006) Synapse formation and function is modulated by the amyloid precursor protein. *J Neurosci* **26**, 7212-7221
149. Turner PR, O'Connor K, Tate WP & Abraham WC (2003) Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. *Prog Neurobiol* **70**, 1-32
150. Kong GK, Adams JJ, Harris HH, Boas JF, Curtain CC, Galatis D, Masters CL, Barnham KJ, McKinstry WJ, Cappai R, et al. (2007) Structural studies of the Alzheimer's amyloid precursor protein copper-binding domain reveal how it binds copper ions. *J Mol Biol* **367**, 148-161
151. Jarrett JT, Berger EP & Lansbury PT, Jr. (1993) The carboxy terminus of the beta amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease. *Biochemistry* **32**, 4693-4697
152. Kang J, Lemaire HG, Unterbeck A, Salbaum JM, Masters CL, Grzeschik KH, Multhaup G, Beyreuther K & Muller-Hill B (1987) The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature* **325**, 733-736
153. Thinakaran G & Koo EH (2008) Amyloid precursor protein trafficking, processing, and function. *J Biol Chem* **283**, 29615-29619
154. Harper JD, Wong SS, Lieber CM & Lansbury PT (1997) Observation of metastable Abeta amyloid protofibrils by atomic force microscopy. *Chem Biol* **4**, 119-125
155. Bitan G, Kirkitadze MD, Lomakin A, Vollers SS, Benedek GB & Teplow DB (2003) Amyloid beta - protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways. *Proc Natl Acad Sci U S A* **100**, 330-335
156. McGowan E, Pickford F, Kim J, Onstead L, Eriksen J, Yu C, Skipper L, Murphy MP, Beard J, Das P, et al. (2005) Abeta42 is essential for parenchymal and vascular amyloid deposition in mice. *Neuron* **47**, 191-199

157. Demuro A, Parker I & Stutzmann GE (2010) Calcium signaling and amyloid toxicity in Alzheimer disease. *J Biol Chem* **285**, 12463-12468
158. Yankner BA, Duffy LK & Kirschner DA (1990) Neurotrophic and neurotoxic effects of amyloid beta protein: reversal by tachykinin neuropeptides. *Science* **250**, 279-282
159. Chen X & Yan SD (2006) Mitochondrial Abeta: a potential cause of metabolic dysfunction in Alzheimer's disease. *IUBMB Life* **58**, 686-694
160. Greig NH, Mattson MP, Perry T, Chan SL, Giordano T, Sambamurti K, Rogers JT, Ovadia H & Lahiri DK (2004) New therapeutic strategies and drug candidates for neurodegenerative diseases: p53 and TNF-alpha inhibitors, and GLP-1 receptor agonists. *Ann N Y Acad Sci* **1035**, 290-315
161. Schmitz C, Rutten BP, Pielen A, Schafer S, Wirths O, Tremp G, Czech C, Blanchard V, Multhaup G, Rezaie P, et al. (2004) Hippocampal neuron loss exceeds amyloid plaque load in a transgenic mouse model of Alzheimer's disease. *Am J Pathol* **164**, 1495-1502
162. Hsia AY, Masliah E, McConlogue L, Yu GQ, Tatsuno G, Hu K, Kholodenko D, Malenka RC, Nicoll RA & Mucke L (1999) Plaque-independent disruption of neural circuits in Alzheimer's disease mouse models. *Proc Natl Acad Sci U S A* **96**, 3228-3233
163. Lue LF, Kuo YM, Roher AE, Brachova L, Shen Y, Sue L, Beach T, Kurth JH, Rydel RE & Rogers J (1999) Soluble amyloid beta peptide concentration as a predictor of synaptic change in Alzheimer's disease. *Am J Pathol* **155**, 853-862
164. McLean CA, Cherny RA, Fraser FW, Fuller SJ, Smith MJ, Beyreuther K, Bush AI & Masters CL (1999) Soluble pool of Abeta amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. *Ann Neurol* **46**, 860-866
165. Walsh DM, Klyubin I, Fadeeva JV, Cullen WK, Anwyl R, Wolfe MS, Rowan MJ & Selkoe DJ (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* **416**, 535-539
166. Lesne S, Koh MT, Kotilinek L, Kayed R, Glabe CG, Yang A, Gallagher M & Ashe KH (2006) A specific amyloid-beta protein assembly in the brain impairs memory. *Nature* **440**, 352-357
167. Barghorn S, Nimmrich V, Striebinger A, Krantz C, Keller P, Janson B, Bahr M, Schmidt M, Bitner RS, Harlan J, et al. (2005) Globular amyloid beta-peptide oligomer - a homogenous and stable neuropathological protein in Alzheimer's disease. *J Neurochem* **95**, 834-847
168. Bernstein SL, Dupuis NF, Lazo ND, Wyttenbach T, Condron MM, Bitan G, Teplow DB, Shea JE, Ruotolo BT, Robinson CV, et al. (2009) Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat Chem* **1**, 326-331
169. Bitan G, Vollers SS & Teplow DB (2003) Elucidation of primary structure elements controlling early amyloid beta-protein oligomerization. *J Biol Chem* **278**, 34882-34889
170. Chen YR & Glabe CG (2006) Distinct early folding and aggregation properties of Alzheimer amyloid-beta peptides Abeta40 and Abeta42: stable trimer or tetramer formation by Abeta42. *J Biol Chem* **281**, 24414-24422
171. Mastrangelo IA, Ahmed M, Sato T, Liu W, Wang C, Hough P & Smith SO (2006) High-resolution atomic force microscopy of soluble Abeta42 oligomers. *J Mol Biol* **358**, 106-119
172. Torok M, Milton S, Kayed R, Wu P, McIntire T, Glabe CG & Langen R (2002) Structural and dynamic features of Alzheimer's Abeta peptide in amyloid fibrils studied by site-directed spin labeling. *J Biol Chem* **277**, 40810-40815
173. Ahmed M, Davis J, Aucoin D, Sato T, Ahuja S, Aimoto S, Elliott JI, Van Nostrand WE & Smith SO (2010) Structural conversion of neurotoxic amyloid-beta(1-42) oligomers to fibrils. *Nat Struct Mol Biol* **17**, 561-567
174. Dong J, Atwood CS, Anderson VE, Siedlak SL, Smith MA, Perry G & Carey PR (2003) Metal binding and oxidation of amyloid-beta within isolated senile plaque cores: Raman microscopic evidence. *Biochemistry* **42**, 2768-2773
175. Grundke-Iqbal I, Fleming J, Tung YC, Lassmann H, Iqbal K & Joshi JG (1990) Ferritin is a component of the neuritic (senile) plaque in Alzheimer dementia. *Acta Neuropathol* **81**, 105-110
176. Quintana C, Bellefqih S, Laval JY, Guerquin-Kern JL, Wu TD, Avila J, Ferrer I, Arranz R & Patino C (2006) Study of the localization of iron, ferritin, and hemosiderin in Alzheimer's disease hippocampus by analytical microscopy at the subcellular level. *J Struct Biol* **153**, 42-54
177. Zecca L, Youdim MB, Riederer P, Connor JR & Crichton RR (2004) Iron, brain ageing and neurodegenerative disorders. *Nature reviews* **5**, 863-873

178. Religa D, Strozyk D, Cherny RA, Volitakis I, Haroutunian V, Winblad B, Naslund J & Bush AI (2006) Elevated cortical zinc in Alzheimer disease. *Neurology* **67**, 69-75
179. Barnham KJ & Bush AI (2008) Metals in Alzheimer's and Parkinson's diseases. *Current opinion in chemical biology* **12**, 222-228
180. Su B, Wang X, Nunomura A, Moreira PI, Lee HG, Perry G, Smith MA & Zhu X (2008) Oxidative stress signaling in Alzheimer's disease. *Curr Alzheimer Res* **5**, 525-532
181. Khan A, Dobson JP & Exley C (2006) Redox cycling of iron by Abeta42. *Free Radic Biol Med* **40**, 557-569
182. Ali FE, Separovic F, Barrow CJ, Cherny RA, Fraser F, Bush AI, Masters CL & Barnham KJ (2005) Methionine regulates copper/hydrogen peroxide oxidation products of Abeta. *J Pept Sci* **11**, 353-360
183. Metodiewa D (1998) Molecular mechanisms of cellular injury produced by neurotoxic amino acids that generate reactive oxygen species. *Amino Acids* **14**, 181-187
184. Maeda J, Ji B, Irie T, Tomiyama T, Maruyama M, Okauchi T, Staufenbiel M, Iwata N, Ono M, Saido TC, et al. (2007) Longitudinal, quantitative assessment of amyloid, neuroinflammation, and anti-amyloid treatment in a living mouse model of Alzheimer's disease enabled by positron emission tomography. *J Neurosci* **27**, 10957-10968
185. Atwood CS, Perry G, Zeng H, Kato Y, Jones WD, Ling KQ, Huang X, Moir RD, Wang D, Sayre LM, et al. (2004) Copper mediates dityrosine cross-linking of Alzheimer's amyloid-beta. *Biochemistry* **43**, 560-568
186. Barnham KJ, Haeffner F, Ciccotosto GD, Curtain CC, Tew D, Mavros C, Beyreuther K, Carrington D, Masters CL, Cherny RA, et al. (2004) Tyrosine gated electron transfer is key to the toxic mechanism of Alzheimer's disease beta-amyloid. *FASEB J* **18**, 1427-1429
187. Danscher G & Stoltenberg M (2005) Zinc-specific autometallographic in vivo selenium methods: tracing of zinc-enriched (ZEN) terminals, ZEN pathways, and pools of zinc ions in a multitude of other ZEN cells. *J Histochem Cytochem* **53**, 141-153
188. Schlief ML, Craig AM & Gitlin JD (2005) NMDA receptor activation mediates copper homeostasis in hippocampal neurons. *J Neurosci* **25**, 239-246
189. Uchida Y, Gomi F, Masumizu T & Miura Y (2002) Growth inhibitory factor prevents neurite extension and the death of cortical neurons caused by high oxygen exposure through hydroxyl radical scavenging. *J Biol Chem* **277**, 32353-32359
190. Rogers JT, Randall JD, Cahill CM, Eder PS, Huang X, Gunshin H, Leiter L, McPhee J, Sarang SS, Utsuki T, et al. (2002) An iron-responsive element type II in the 5'-untranslated region of the Alzheimer's amyloid precursor protein transcript. *J Biol Chem* **277**, 45518-45528
191. Son M, Puttaparthi K, Kawamata H, Rajendran B, Boyer PJ, Manfredi G & Elliott JL (2007) Overexpression of CCS in G93A-SOD1 mice leads to accelerated neurological deficits with severe mitochondrial pathology. *Proc Natl Acad Sci U S A* **104**, 6072-6077
192. Lee JY, Cole TB, Palmiter RD, Suh SW & Koh JY (2002) Contribution by synaptic zinc to the gender-disparate plaque formation in human Swedish mutant APP transgenic mice. *Proc Natl Acad Sci U S A* **99**, 7705-7710
193. Rovira C, Arbez N & Mariani J (2002) Abeta(25-35) and Abeta(1-40) act on different calcium channels in CA1 hippocampal neurons. *Biochem Biophys Res Commun* **296**, 1317-1321
194. McLaurin J & Chakrabartty A (1996) Membrane disruption by Alzheimer beta-amyloid peptides mediated through specific binding to either phospholipids or gangliosides. Implications for neurotoxicity. *J Biol Chem* **271**, 26482-26489
195. Inoue S (2008) In situ Abeta pores in AD brain are cylindrical assembly of Abeta protofilaments. *Amyloid* **15**, 223-233
196. Isaacs AM, Senn DB, Yuan M, Shine JP & Yankner BA (2006) Acceleration of amyloid beta-peptide aggregation by physiological concentrations of calcium. *J Biol Chem* **281**, 27916-27923
197. Mattson MP (2007) Calcium and neurodegeneration. *Aging Cell* **6**, 337-350
198. Bertoncini CW, Jung YS, Fernandez CO, Hoyer W, Griesinger C, Jovin TM & Zweckstetter M (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc Natl Acad Sci U S A* **102**, 1430-1435
199. Lee JC, Gray HB & Winkler JR (2005) Tertiary contact formation in alpha-synuclein probed by electron transfer. *J Am Chem Soc* **127**, 16388-16389

200. Lee JC, Langen R, Hummel PA, Gray HB & Winkler JR (2004) Alpha-synuclein structures from fluorescence energy-transfer kinetics: implications for the role of the protein in Parkinson's disease. *Proc Natl Acad Sci U S A* **101**, 16466-16471
201. Tamamizu-Kato S, Kosaraju MG, Kato H, Raussens V, Ruysschaert JM & Narayanaswami V (2006) Calcium-triggered membrane interaction of the alpha-synuclein acidic tail. *Biochemistry* **45**, 10947-10956
202. Nielsen MS, Vorum H, Lindersson E & Jensen PH (2001) Ca2+ binding to alpha-synuclein regulates ligand binding and oligomerization. *J Biol Chem* **276**, 22680-22684
203. Uversky VN, Lee HJ, Li J, Fink AL & Lee SJ (2001) Stabilization of partially folded conformation during alpha-synuclein oligomerization in both purified and cytosolic preparations. *J Biol Chem* **276**, 43495-43498
204. Fernandez CO, Hoyer W, Zweckstetter M, Jares-Erijman EA, Subramaniam V, Griesinger C & Jovin TM (2004) NMR of alpha-synuclein-polyamine complexes elucidates the mechanism and kinetics of induced aggregation. *EMBO J* **23**, 2039-2046
205. Dexter DT, Carayon A, Javoy-Agid F, Agid Y, Wells FR, Daniel SE, Lees AJ, Jenner P & Marsden CD (1991) Alterations in the levels of iron, ferritin and other trace metals in Parkinson's disease and other neurodegenerative diseases affecting the basal ganglia. *Brain* **114 (Pt 4)**, 1953-1975
206. Riederer P, Sofic E, Rausch WD, Schmidt B, Reynolds GP, Jellinger K & Youdim MB (1989) Transition metals, ferritin, glutathione, and ascorbic acid in parkinsonian brains. *J Neurochem* **52**, 515-520
207. Hirsch EC, Brandel JP, Galle P, Javoy-Agid F & Agid Y (1991) Iron and aluminum increase in the substantia nigra of patients with Parkinson's disease: an X-ray microanalysis. *J Neurochem* **56**, 446-451
208. Berg D, Gerlach M, Youdim MB, Double KL, Zecca L, Riederer P & Becker G (2001) Brain iron pathways and their relevance to Parkinson's disease. *J Neurochem* **79**, 225-236
209. Oestreicher E, Sengstock GJ, Riederer P, Olanow CW, Dunn AJ & Arendash GW (1994) Degeneration of nigrostriatal dopaminergic neurons increases iron within the substantia nigra: a histochemical and neurochemical study. *Brain Res* **660**, 8-18
210. Gotz ME, Double K, Gerlach M, Youdim MB & Riederer P (2004) The relevance of iron in the pathogenesis of Parkinson's disease. *Ann N Y Acad Sci* **1012**, 193-208
211. Paik SR, Shin HJ, Lee JH, Chang CS & Kim J (1999) Copper(II)-induced self-oligomerization of alpha-synuclein. *The Biochemical journal* **340 (Pt 3)**, 821-828
212. Khan A, Ashcroft AE, Higenell V, Korchazhkina OV & Exley C (2005) Metals accelerate the formation and direct the structure of amyloid fibrils of NAC. *J Inorg Biochem* **99**, 1920-1927
213. Paik SR, Lee JH, Kim DH, Chang CS & Kim J (1997) Aluminum-induced structural alterations of the precursor of the non-A beta component of Alzheimer's disease amyloid. *Arch Biochem Biophys* **344**, 325-334
214. Hoyer W, Antony T, Cherny D, Heim G, Jovin TM & Subramaniam V (2002) Dependence of alpha-synuclein aggregate morphology on solution conditions. *J Mol Biol* **322**, 383-393
215. Golts N, Snyder H, Frasier M, Theisler C, Choi P & Wolozin B (2002) Magnesium inhibits spontaneous and iron-induced aggregation of alpha-synuclein. *J Biol Chem* **277**, 16116-16123
216. Liu LL & Franz KJ (2007) Phosphorylation-dependent metal binding by alpha-synuclein peptide fragments. *J Biol Inorg Chem* **12**, 234-247
217. Binolfi A, Rasia RM, Bertoncini CW, Ceolin M, Zweckstetter M, Griesinger C, Jovin TM & Fernandez CO (2006) Interaction of alpha-synuclein with divalent metal ions reveals key differences: a link between structure, binding specificity and fibrillation enhancement. *J Am Chem Soc* **128**, 9893-9901
218. Bharathi & Rao KS (2007) Thermodynamics imprinting reveals differential binding of metals to alpha-synuclein: relevance to Parkinson's disease. *Biochem Biophys Res Commun* **359**, 115-120
219. Conway KA, Lee SJ, Rochet JC, Ding TT, Harper JD, Williamson RE & Lansbury PT, Jr. (2000) Accelerated oligomerization by Parkinson's disease linked alpha-synuclein mutants. *Ann N Y Acad Sci* **920**, 42-45
220. Conway KA, Lee SJ, Rochet JC, Ding TT, Williamson RE & Lansbury PT, Jr. (2000) Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy. *Proc Natl Acad Sci U S A* **97**, 571-576

221. Lowe R, Pountney DL, Jensen PH, Gai WP & Voelcker NH (2004) Calcium(II) selectively induces alpha-synuclein annular oligomers via interaction with the C-terminal domain. *Protein Sci* **13**, 3245-3252
222. Adamczyk A & Strosznajder JB (2006) Alpha-synuclein potentiates Ca2+ influx through voltage-dependent Ca2+ channels. *Neuroreport* **17**, 1883-1886
223. Schulz JB (2007) Mechanisms of neurodegeneration in idiopathic Parkinson's disease. *Parkinsonism Relat Disord* **13 Suppl 3**, S306-308
224. Vassallo N & Herms J (2003) Cellular prion protein function in copper homeostasis and redox signalling at the synapse. *J Neurochem* **86**, 538-544
225. Detwiler LA (1992) Scrapie. *Rev Sci Tech* **11**, 491-537
226. Prusiner SB (1998) Prions. *Proc Natl Acad Sci U S A* **95**, 13363-13383
227. Wuthrich K & Riek R (2001) Three-dimensional structures of prion proteins. *Adv Protein Chem* **57**, 55-82
228. Whittal RM, Ball HL, Cohen FE, Burlingame AL, Prusiner SB & Baldwin MA (2000) Copper binding to octarepeat peptides of the prion protein monitored by mass spectrometry. *Protein Sci* **9**, 332-343
229. Kramer ML, Kratzin HD, Schmidt B, Romer A, Windl O, Liemann S, Hornemann S & Kretzschmar H (2001) Prion protein binds copper within the physiological concentration range. *J Biol Chem* **276**, 16711-16719
230. Que EL, Domaille DW & Chang CJ (2008) Metals in neurobiology: probing their chemistry and biology with molecular imaging. *Chem Rev* **108**, 1517-1549
231. Brown DR, Qin K, Herms JW, Madlung A, Manson J, Strome R, Fraser PE, Kruck T, von Bohlen A, Schulz-Schaeffer W, et al. (1997) The cellular prion protein binds copper in vivo. *Nature* **390**, 684-687
232. Walter ED, Stevens DJ, Visconte MP & Millhauser GL (2007) The prion protein is a combined zinc and copper binding protein: Zn2+ alters the distribution of Cu2+ coordination modes. *J Am Chem Soc* **129**, 15440-15441
233. Choi CJ, Kanthasamy A, Anantharam V & Kanthasamy AG (2006) Interaction of metals with prion protein: possible role of divalent cations in the pathogenesis of prion diseases. *Neurotoxicology* **27**, 777-787
234. Jackson GS, Murray I, Hosszu LL, Gibbs N, Waltho JP, Clarke AR & Collinge J (2001) Location and properties of metal-binding sites on the human prion protein. *Proc Natl Acad Sci U S A* **98**, 8531-8535
235. Yam AY, Gao CM, Wang X, Wu P & Peretz D (2010) The octarepeat region of the prion protein is conformationally altered in PrP(Sc). *PLoS One* **5**, e9316
236. Fernaeus S & Land T (2005) Increased iron-induced oxidative stress and toxicity in scrapie-infected neuroblastoma cells. *Neurosci Lett* **382**, 217-220
237. Fernaeus S, Reis K, Bedecs K & Land T (2005) Increased susceptibility to oxidative stress in scrapie-infected neuroblastoma cells is associated with intracellular iron status. *Neurosci Lett* **389**, 133-136
238. Giese A, Buchholz M, Herms J & Kretzschmar HA (2005) Mouse brain synaptosomes accumulate copper-67 efficiently by two distinct processes independent of cellular prion protein. *J Mol Neurosci* **27**, 347-354
239. Quaglio E, Chiesa R & Harris DA (2001) Copper converts the cellular prion protein into a protease-resistant species that is distinct from the scrapie isoform. *J Biol Chem* **276**, 11432-11438
240. Bocharova OV, Breydo L, Salnikov VV & Baskakov IV (2005) Copper(II) inhibits in vitro conversion of prion protein into amyloid fibrils. *Biochemistry* **44**, 6776-6787
241. Hijazi N, Shaked Y, Rosenmann H, Ben-Hur T & Gabizon R (2003) Copper binding to PrPC may inhibit prion disease propagation. *Brain Res* **993**, 192-200
242. Sigurdsson EM, Brown DR, Alim MA, Scholtzova H, Carp R, Meeker HC, Prelli F, Frangione B & Wisniewski T (2003) Copper chelation delays the onset of prion disease. *J Biol Chem* **278**, 46199-46202
243. Brown DR, Clive C & Haswell SJ (2001) Antioxidant activity related to copper binding of native prion protein. *J Neurochem* **76**, 69-76
244. Brown DR, Wong BS, Hafiz F, Clive C, Haswell SJ & Jones IM (1999) Normal prion protein has an activity like that of superoxide dismutase. *The Biochemical journal* **344 Pt 1**, 1-5
245. Nishina K, Deleault NR, Lucassen RW & Supattapone S (2004) In vitro prion protein conversion in detergent-solubilized membranes. *Biochemistry* **43**, 2613-2621

246. Kim NH, Choi JK, Jeong BH, Kim JI, Kwon MS, Carp RI & Kim YS (2005) Effect of transition metals (Mn, Cu, Fe) and deoxycholic acid (DA) on the conversion of PrPC to PrPres. *FASEB J* **19**, 783-785

247. Levin J, Bertsch U, Kretzschmar H & Giese A (2005) Single particle analysis of manganese-induced prion protein aggregates. *Biochem Biophys Res Commun* **329**, 1200-1207

248. Brown DR, Hafiz F, Glasssmith LL, Wong BS, Jones IM, Clive C & Haswell SJ (2000) Consequences of manganese replacement of copper for prion protein function and proteinase resistance. *EMBO J* **19**, 1180-1186

7. The S100 protein family

Part of this chapter was published in

Fritz G, Botelho HM, Morozova-Roche LA, Gomes CM (2010) Natural and amyloid self-assembly of S100 proteins: structural basis of functional diversity. *FEBS J*, 277(22) 4578-4590

7.1. Family description

The S100 protein family represents the largest subgroup within the Ca^{2+}-binding EF-hand superfamily. They are only found in vertebrates. The name of the protein family has derived from the fact that the first identified S100 proteins were obtained from the soluble (S) bovine brain fraction upon fractionation with saturated (100%) ammonium sulfate [1]. The genes encoding the large majority of human S100 proteins are organized in a gene cluster located in chromosomal region 1q21 [2-3]. This region harbors the genes of S100A1 to S100A16, which are the result of several gene duplication events. The genes of other S100 proteins like S100B, S100P or S100Z are located in humans in chromosomes 21, 4 and 5, respectively.

In humans, 21 different S100 proteins are identified up to date (Figure 7.1) and similar numbers have been found in other mammalia based on genomic analysis. Further diverse branches of S100 proteins were found in other vertebrates. The level of sequence identity among the S100 proteins within one species varies considerably, *e. g.* for human proteins the identity ranges between 22% and 57%. Noteworthy, many S100 proteins exhibit very distinctive expression patterns in different tissues and cell types, as well as specific subcellular localization, underlining the high degree of specialization among them. Corresponding to their diversity in primary structure and localization, the S100 proteins are involved in the regulation of a multitude of

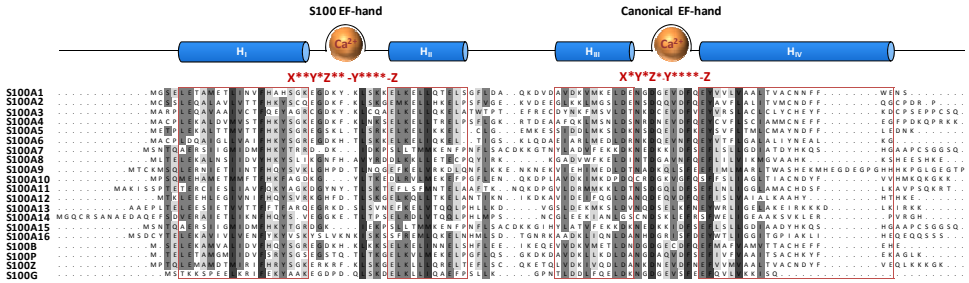


Figure 7.1 – Sequence alignment of human S100 proteins. The sketch depicts the consensus location of helices (cylinders) and random coil (line). Red letters indicate the calcium ligands.

cellular processes like cell cycle control, cell growth, differentiation and motility. Considering the diverse S100 protein functions, it is no surprise to find that these proteins are implicated in numerous human diseases, like different types of cancers characterized by altered expression levels of S100 proteins [4], neurodegenerative disorders such as Alzheimer's disease [5-6], inflammatory and autoimmune diseases [4].

The conformational properties and function of S100 proteins are modulated by metal ion binding. Most S100 proteins are Ca^{2+} signal transducers. The binding of Ca^{2+} to EF-hand type domains triggers conformational changes allowing interactions with other proteins. The exception is S100G, which acts as a Ca^{2+} buffer, sequestering free Ca^{2+} in the cytoplasm after a concentration raise without significant conformational changes [7]. In many S100 proteins, additional binding of Zn^{2+} fine tunes protein folding and function [8-9]. Intracellularly, S100 proteins act as Ca^{2+} sensors, translating intracellular Ca^{2+} level increase into a cellular response. An increasing number of S100 proteins is also reported to occur extracellularly binding to the Receptor for Advanced Glycation Endproducts (RAGE) [10-13] or Toll-like receptor 4 [14]. Recently, a new property among S100 proteins was unveiled: S100A8/A9 proteins can form amyloids in a metal ion mediated fibrillization process in the ageing prostate [15]. In the following sections these aspects and the possible functional and biological implications of physiological amyloid formation by S100 proteins will be addressed.

7.2. Structural properties of S100 proteins

Most S100 protein family members form homo- and heterodimers, but with largely different preferences. Recently, it was shown that some S100 proteins might exist as monomers at very low concentrations in the cell, although the physiological function of the monomers is unclear [16]. It was proposed that equilibrium between monomers and homodimers might

facilitate the formation of heterodimers in the cell [16-17]. Several heterodimeric S100 proteins have been reported, but only the S100A8/A9 heterodimer is well characterized [14, 18-21]. The list of S100 heterodimers is steadily growing: S100B forms heterodimers with S100A1 [22], S100A6 [23-24] and S100A11 [24]; S100A1 with S100A4 [25] and S100P [26]; and S100A7 with S100A10 [27]. Non-covalent multimers were observed for S100A12 [28], S100A8/A9 [18, 29], S100B [12], S100A4 [30] and a Zn^{2+}-dependent tetramer for S100A2 [31]. Comparison of the structure of S100A8/A9 with those of the corresponding homodimers revealed that the solvent exposed area is reduced in the heterodimer, which might represent the driving force of heterodimer formation [18]. It is proposed that heterodimer formation apart from homodimeric assembly might lead to further diversification of S100 protein functions [17, 32].

7.2.1. EF-hand Ca^{2+} binding

To date there is ample 3D structural information available for individual S100 proteins in both the inactive Ca^{2+}-free and active Ca^{2+}-loaded states [33]. Despite the variation in amino acid sequence, the 3D structures of S100 proteins exhibit key structural features common to all members of the family. Each S100 monomer is about 10-12 kDa in size and composed of two EF-hand helix-loop-helix structural motifs, which are arranged in a back-to-back manner and connected by a flexible linker. The C-terminal EF-hand contains the classical Ca^{2+}-binding motif, common to all EF-hand proteins. The loop has a typical sequence signature of 12 amino acids flanked by helices H_{III} and H_{IV} (Figure 7.2B). Ca^{2+} is coordinated mainly by amino acid side chains. The N-terminal EF-hand exhibits a slightly different architecture and contains a specific 14 amino acid motif flanked by helices H_I and H_{II} (Figure 7.2A) where Ca^{2+} is mainly coordinated by backbone carbonyls (Figure 7.3) [33]. This motif is characteristic for S100 proteins and therefore it is often called "S100-specific" or "pseudo EF-hand". In both EF-hands, Ca^{2+} is

heptacoordinated in a pentagonal bipyramidal geometry. In S100 proteins, the ligands are denoted by X, Y, Z, –Y and –Z (Figure 7.3) [33-34]. The position –X is occupied by a water molecule. The glutamate or aspartate residue at position –Z binds Ca^{2+} in a bidentate manner. It is the displacement of this amino acid residue that determines the EF-hand opening and Ca^{2+}/Mg^{2+} selectivity [35]. The S100 protein dimer interface is formed by helices H_I and H_{IV} from both monomers, building a compact four helix bundle as core structure of the S100 protein dimer (Figure 7.2C-D).

Typically, S100 proteins bind two Ca^{2+} ions per monomer (K_d = 20-500 μM) with strong positive cooperativity [33]. In the absence of Ca^{2+}, EF-hands

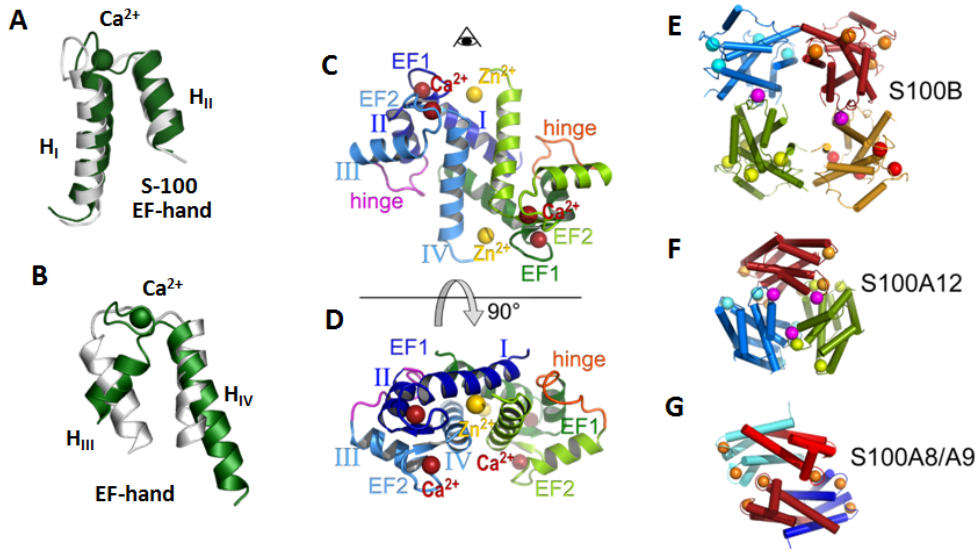


Figure 7.2 - Structure of S100 Proteins. (A-B) Calcium-driven conformational changes at the EF-hands in S100 proteins. Structure of the N-terminal, S100-specific EF-hand (A) and the C-terminal, canonical EF-hand (B) in the metal-free (lighter) and Ca^{2+}-bound (darker) form of S100A6. The EF-hand flanking helices (H_I-H_{IV}) are identified. (C-D) Structure of the human S100B homodimer loaded with Ca^{2+} and Zn^{2+} (Ostendorp, T., Diez, J., Heizmann, C.W., Fritz, G., unpublished work, 3D10). (C) Side view; (D) Top view. The monomers are shown in blue and green respectively. The N-terminal S100 specific EF-hand (EF-hand 1) is shown in dark color, the C-terminal canonical EF-hand in brighter color (EF-hand 2). The hinges connecting both EF-hands are shown in magenta and orange. The four bound Ca^{2+} ions are shown as red spheres. The two Zn^{2+} bound at the dimer interface of S100B are shown as yellow spheres. (E-G) Multimeric states of S100 proteins. S100B octamer, 2H61 (E), S100A12 hexamer, 1GQM (F) and S100A8/A9 tetramer, 1XK4 (G). Each dimer in S100B or S100A12 is shown in individual color. S100A8 is shown in red, S100A9 in blue. Bound Ca^{2+} ions are shown as spheres; intersubunit Ca^{2+} ions are shown as magenta spheres.

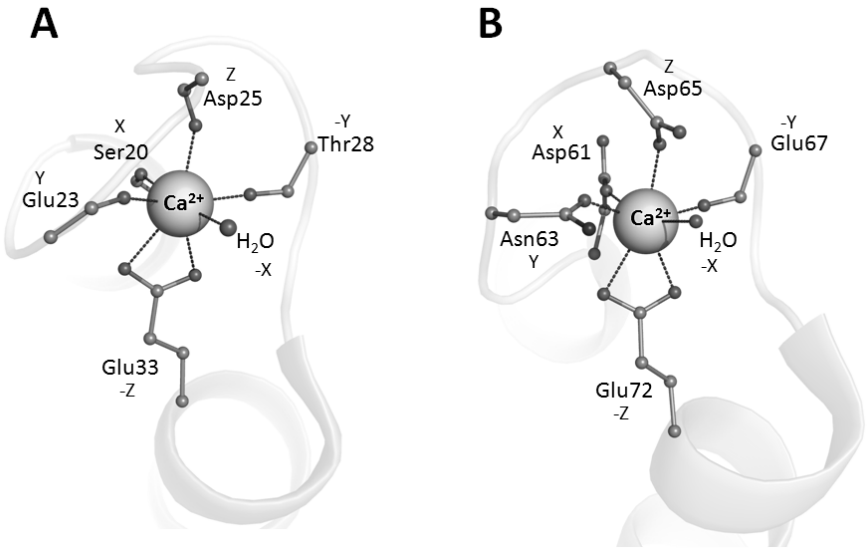


Figure 7.3 – Ca^{2+}-bound EF-hands in S100A6. (A) N-terminal S100-specific EF-hand (B) C-terminal canonical EF-hand. Ca^{2+}-coordinating residues are shown as lines and bead. Coordination bonds are depicted as dashed lines. From [33].

are occupied by water [36-37], Na^{+} [38] or Mg^{2+}. The much lower Mg^{2+} affinity (K_d = 1-125 mM) [39-40] does not interfere with Ca^{2+} binding at physiological Mg^{2+} concentrations (0.5 mM) [39, 41-43]. Due to amino acid substitutions in the critical Z position, the EF-hands in S100A16 are not able to bind Ca^{2+} [44].

7.2.2. Zn^{2+} binding sites

Many S100 proteins are reported to bind Zn^{2+} with high affinity (K_d = 4 nM to 2 mM) [9]. The Zn^{2+}-binding S100 proteins can be subdivided into two subgroups. One group involves Cys residues in Zn^{2+} coordination, whereas the second group binds Zn^{2+} exclusively via the side-chains of His, Glu and Asp residues. The first group has been characterized so far by spectroscopic analysis in combination with molecular modeling, showing *e. g.* for S100A2 that Zn^{2+} is coordinated by residues from different monomers [31]. For the second group, encompassing S100A7, S100A8/A9, S100A12 and S100B, detailed structural information mainly by X-ray crystallography is available.

S100A7, S100A12 and S100B bind two Zn^{2+} ions per homodimer at the dimer interface. The Zn^{2+} ions are coordinated by residues originating from both subunits, thereby additionally stabilizing the dimer [45-47].

7.2.3. Cu^{2+} binding sites

Some S100 proteins like S100A5 [48], S100A12 [49], S100A13 [50] and S100B [51] also bind Cu^{2+}. This happens in the same binding sites as Zn^{2+} but with different affinity (K_d = 0.46 – 55 μM) [48, 51-52].

7.3. Metal ions as modulators of S100 conformation and stability

The metal binding properties of S100 proteins have a pivotal influence as modulators of their conformation, folding, oligomerization state and ultimately, function. As outlined above, S100 proteins are able to bind different metal ions, among Ca^{2+}, Zn^{2+} and Cu^{2+}. Generally, the dimeric S100 proteins bind four Ca^{2+} ions per dimer with micromolar to hundreds micromolar binding constants and strong cooperativity. In the Ca^{2+}-free state the helices of both EF-hands in each monomer adopt an antiparallel conformation masking the target protein interaction site. Upon Ca^{2+} binding, the C-terminus undergoes a major conformational change involving motions of the two helices which flank the Ca^{2+} binding residues (Figure 7.2B). Helix H_{III} of the canonical EF-hand makes a 90° movement, opening the structure, whereas the N-terminal EF-hand exhibits only minor structural changes (Figure 7.2A-B). This Ca^{2+}-induced structural change leads to the exposure of a wide hydrophobic cleft which mediates target recognition. This surface is formed by residues of the hinge region, helix H_{III}, and the C-terminus, the regions exhibiting the largest variation in amino acid sequence throughout the S100 family. The dimer interface is built by residues of helices H_I and H_{IV} and is well conserved across the different S100 proteins. Also, these helices barely move during Ca^{2+} binding, maintaining the dimeric state of the S100 proteins during signaling. The invariability of the residues and the conserved

spatial arrangement of the helices at the dimer interface are the basis for the formation of the different observed heterodimers.

Structural studies on apo S100A2 have shown that in the absence of Ca^{2+}, a Na^{+} ion resides in the N-terminal S100-specific EF-hand [38]. In contrast to the classical EF-hand that coordinates Ca^{2+} through four acidic side chains originating from Asn, Asp or Glu residues, the N-terminal EF-hand provides only one strictly conserved Glu side chain for Ca^{2+} coordination. Therefore, the Na+ ion can compensate partially for the negative charge in the Ca^{2+}-free state. Ca^{2+} binding to the N-terminal EF-hand causes a movement of this Glu residue towards the bound ion, pulling helix H_{II} slightly towards the Ca^{2+} site. Similar to many other EF-hand proteins, S100 proteins bind also Mg^{2+} ions into their EF-hand sites but the reported affinities for Mg^{2+} ions are rather low having only a minor effect on Ca^{2+} binding.

Next to Ca^{2+}, many S100 proteins (S100B, S100A2, S100A3, S100A6, S100A7, S100A8/9, S100A12) bind Zn^{2+} in specific sites, whose metallation state also influences protein conformation, folding and presumably function. Binding of Zn^{2+} occurs invariably at interfaces, involving coordinating residues from different subunits that are close to the dimer interface. One of these proteins is S100A7, that is upregulated in the keratinocytes of patients suffering from the chronic skin disease psoriasis, and that has been hypothesized to account for the microbial resistance of skin [53]. The structure of this protein has elicited two identical high affinity Zn^{2+} binding sites formed by His/Asp residues from different monomers that “clip” together the two subunits. A substantial stabilization of the dimer is expected to arise from Zn^{2+} binding, as it promotes head-to-tail interactions between the two monomers, although in this particular case Zn^{2+} does not seem to be essential for protein stability [46].

There is evidence for an interesting cross talk between Ca^{2+} and Zn^{2+} binding to S100 proteins, illustrating how binding of different metal ions

results in conformational adjustments and modulation of protein folding and function. In S100B and S100A12 Zn^{2+} binding leads to an increase of the Ca^{2+} affinity [54-55], whereas in S100A2 the opposite effect was observed, *i. e.* Zn^{2+} decreased Ca^{2+} affinity pointing to interplay of the metal ions in the activation of S100 proteins [31]. For S100A12 and S100B, the molecular mechanism of increase of Ca^{2+} affinity by Zn^{2+} can be deduced from the structural information available (Ostendorp, T., Diez, J., Heizmann, C.W., Fritz, G., unpublished work) [45]. In both proteins there is one Zn^{2+}-coordinating His residue located in the Ca^{2+}-binding loop, which might help to stabilize the Ca^{2+}bound conformation, thereby increasing Ca^{2+} affinity. The structure of S100A12 with only bound Zn^{2+} shows also that Zn^{2+} alone already can induce structural changes similar to those induced by Ca^{2+}, which will also lead to an increase in Ca^{2+} affinity. Other Zn^{2+} coordinating residues are located in the C-terminus of the S100 proteins. Zn^{2+} coordination leads to a stabilization and extension of the of the C-terminal helix, changing the orientation of residues involved in target binding. As expected from these structural changes, Zn^{2+} binding modulates target binding properties of the different S100 proteins. *e. g.*, Zn^{2+} prevents S100A8/A9 binding to arachidonic acid [56]. On the other hand, Zn^{2+} and Ca^{2+} binding to S100A9 are both required for interaction with receptors like RAGE or Toll-like receptor 4 [10, 14]. Similarly, Zn^{2+} increased the Ca^{2+} dependent interaction of S100A12 with RAGE [57]. In the case of S100B, Zn^{2+} alone already could trigger binding to tau [58-59], or IQGAP1 [60]. Moreover, Zn^{2+} binding enhanced Ca^{2+} dependent interaction with the AHNAK [61] or the target protein derived peptide TRTK-12 [62].

Recent work on the S100A2 protein, a cell cycle regulator which binds and activates p53 in a Ca^{2+} dependent manner, has shown that metal ion binding influences the protein conformation and stability [8]. S100A2 binds two Ca^{2+} and two Zn^{2+} ions per subunit, known to be associated with activation (Ca^{2+})

or inhibition (Zn^{2+}) of downstream signaling. Zn^{2+} binds at distinct sites which have different metal binding affinities, and physiologically relevant Zn^{2+} concentrations decrease the affinity for Ca^{2+} binding, resulting in a blockage of p53 activation. It has been recently elicited that the S100A2 conformation is sensitive to the metallation state, although rearrangements resulting from metal binding preserve the overall fold of the protein: S100A2 is destabilized by Zn^{2+} and stabilized by Ca^{2+}, suggesting a synergistic effect between binding of the different metals. Thus, the decrease in Ca^{2+} affinity through Zn^{2+} is presumably a result of the general destabilization of the protein. Further contributions might come from the exposure of a hydrophobic surface on Zn^{2+} binding, making additional exposure of the hydrophobic surface induced by Ca^{2+} less favorable. The antagonistic effect of Zn^{2+} and Ca^{2+} in the control of S100A2 stability provides a molecular rationale for the action of both metal ions: hypothetically, in tissues expressing S100A2, the Zn^{2+} imbalance, which may arise in some types of cancers as a result of the upregulation of Zn^{2+} transporters [63-64], may contribute to enhanced cell proliferation through destabilization of S100A2. That would impair the interaction with p53 and disrupt subsequent downstream cell cycle regulation. This further illustrates how binding of different metal ions to S100 proteins has the potential to result in conformational adjustments and modulation of protein folding and functions.

A number of S100 proteins also bind Cu^{2+} (S100B [51], S100A5 [48], S100A12 [49] and S100A13 [50]) and this frequently occurs at the same sites to which Zn^{2+} binds. That is for example the case of S100A12, an important protein in the inflammatory response and a factor in host/parasite defenses, which binds Cu^{2+} and Zn^{2+} at the same site and corresponds to the Zn^{2+} binding site in S100A7, evoking a possible similar structural and functional role. S100B, one of the most abundant proteins in the human brain also binds Cu^{2+}, and in this case a putative neuroprotective role was suggested.

7.4. S100 functional oligomers

Metal ions play also a crucial role in the formation of larger oligomeric species of S100 proteins, namely tetramers, hexamers and octamers. These are in many cases essential for biological function and signaling: tetrameric S100B [12] and hexameric S100A12 [65] bind RAGE with higher affinity than the dimeric counterparts, only multimeric S100A4 promotes neurite outgrowth [66], and microtubule formation is only promoted by the Ca^{2+} induced S100A8/A9 tetramer [18]. Ca^{2+}-loaded S100A12 forms a functional hexamer whose quaternary structure is maintained by additional interdimer bridging Ca^{2+} ions, which are coordinated by residues from the C-terminal EF-hand and helix H_{III} from two adjacent dimers. This arrangement of ligands for the interdimer Ca^{2+} "crosslinker" is only possible when the C-terminal EF-hand is in the Ca^{2+} bound state [49]. Similarly, two S100A8/A9 heterodimers can assemble into a heterotetramer in a strictly Ca^{2+}-dependent manner [18]. However, the initial S100A8/A9 heterodimer can be formed in the presence or absence of Ca^{2+}. By contrast, the formation of S100B tetramers is not dependent on Ca^{2+} and the tetramer remains stable in the absence of the metal ion [12]. This difference may result from the additional hydrophobic moieties found in interfaces of S100B, which are essentially polar in S100A12 and S100A8/A9 [12]. Nevertheless, the presence of Ca^{2+} enhances the oligomerization of S100B into hexamers and octamers, and the octameric crystal structure reveals intersubunit Ca^{2+} ions. The oligomerization role is not restricted to Ca^{2+}, as in S100A2 binding of Zn^{2+} to the low affinity site triggers formation of a tetramer via the assembly of two S100A2 dimers [31]. Altogether these results point to a very clear role of metal ions in the formation of functional S100 oligomers. However, novel roles for non-functional S100 oligomers are emerging with the recent finding of metal-dependent amyloid formation processes by S100A8/A9, which will be addressed further.

7.5. Functional diversity of S100 proteins

To date a great number of distinct functions have been attributed to S100 proteins in both the intra- and extracellular milieu. Though S100 proteins appear to lack enzymatic activity themselves, they play biological roles through binding to other proteins and changing the activity of their targets. As we discussed above, the conformation and even oligomerization state of S100s are responsive to Ca^{2+} and consequently they mediate Ca^{2+} signals by binding to other intracellular target proteins and modulating their conformation and activity in a Ca^{2+}- and possibly also in a Zn^{2+}- and Cu^{2+}-dependent manner. Indeed, the assembly into multiple complexes is considered in general as a significant generic mechanism of protein functional diversification via varying their conformational states and associated ligands [67]. Several S100 proteins exhibit Ca^{2+}-dependent interactions with metabolic enzymes (S100A1 and S100B with aldolase C) [68], with kinases (S100B with Ndr or Src kinases) [69-70], with cytoskeletal proteins (S100A1 with tubulin, S100B with CapZ and S100P with ezrin) [71-76] or with DNA binding proteins (S100A2, S100A4 and S100B interact with p53) [77-79]. As a result, intracellularly S100 proteins are involved in the regulation of cell cycle, cell growth and differentiation, apoptosis, migration, calcium homeostasis, protein phosphorylation, cellular motility and other important processes.

Some S100 proteins including S100A4, S100A7, S100A8/A9, S100A11, S100A12, S100B and others can be secreted, exhibiting cytokine-like and chemotactic activity. When S100A7, S100A8, S100A9, S100A12 or S100B are secreted in response to cell damage or activation, they become danger signals, activating other immune and endothelial cells. Accordingly, they were defined as damage-associated molecular pattern (DAMP) molecules in innate immunity [80-81]. The S100A8/A9 complex accounts for up to 40% of total cytosolic proteins in neutrophils and secreted S100A8/A9 as well as

S100A12 are found at high concentrations in inflamed tissues, producing strong pro-inflammatory effects.

S100B is highly expressed in the human brain and actively secreted by astrocytes, neurons, microglia, glioblastoma or Schwann cells [82]. Its extracellular concentration reaches micromolar levels after traumatic brain injury and in neurodegenerative disorders such as Alzheimer's disease or Down's syndrome. The action of S100B is strongly dependent on its concentration: at nanomolar levels it is neuroprotective, while at the micromolar concentration range it promotes apoptosis [83]. Both trophic and toxic effects of extracellular S100B are mediated by RAGE, the multi-ligand Receptor for Advanced Glycation Endproducts [21].

A large number of S100 proteins have been shown to interact with RAGE, including S100A1, S100A2, S100A4, S100A5, S100A6, 100A7, S100A8/A9, S100A11, S100A12 and S100B [20]. For example, S100A8/A9 interacts with a subpopulation of RAGE carrying carboxylated glycans and triggers RAGE-dependent NF-κB activation and cellular proliferation [84]. However the interaction with RAGE does not exclude that S100A8/A9 can act via different mechanisms. It was reported that S100A8/A9 proteins exert apoptosis-inducing activity causing cell death via the mitochondrial pathway [85] as well as that the S100A8/A9-induced cell death occurs via apoptosis and autophagy-like mechanisms including ROS-mediated cross-talk between mitochondria and lysosomes: increasing ROS production by mitochondria is followed by mitochondrial damage and lysosomal activation [86]. In addition, Viemann et al. [87] have reported that S100A8/A9 induce cell death via caspase-dependent and caspase-independent programmes which have features of both apoptosis and necrosis.

S100A8 and S100A9 were found to activate also Toll-like receptor 4, acting as innate amplifiers of inflammation and cancer [88-89]. Specifically, it was suggested that S100A8/A9, Toll-like receptor 4 and SAA3, an important

downstream target for S100A8/A9, represent a paracrine positive-feedback cascade, which is critical for inducing an inflammatory response in pre-metastatic tissues and thus, rendering them susceptible for tumor cell immigration and metastasis formation [89]. Recently it was demonstrated in a mouse model that via activation of Toll-like receptor 4 S100A8 and S100A9 induce the development of systemic autoimmunity [88].

Due to their deregulated expression, response to stress and association with neoplastic, degenerative and autoimmune disorders, S100 proteins gain significant interest as potential therapeutic targets. In a view of the large number of tertiary and quaternary structures adopted by S100s and the complex structure-functional relationship affecting their interactions with the target proteins, it is tempting to speculate that this variability may account for the promiscuity of S100 proteins. Therefore, systematic studies of the conformational changes and oligomerization of S100 proteins will be of critical importance in the development of potential therapeutics.

7.6. Amyloid formation by S100A8/A9 proteins

Recently we have found a new amyloidogenic property of S100A8/A9 proteins, implicating them in another degenerative process in the ageing prostate, specifically, in amyloid formation [15]. The conversion of functional proteins and peptides into insoluble amyloid structures and their deposition in a variety of tissues and organs is a hallmark of a growing number of age-related degenerative disorders, including Alzheimer's and Parkinson's diseases, type II diabetes and systemic amyloidoses. Prostate amyloid deposits known as *corpora amylacea* belong to the type of localized amyloidoses. They are associated with age-related prostate tissue remodeling and occur frequently in middle-aged and elderly men.

These inclusions can vary in size from sub-millimeter to a few millimeters diameter (Figure 7.4A) and can in some instances constitute up to a third of the prostate gland bulk weight. Despite of their high prevalence in later life

[90], their role in prostate benign and malignant changes is still disputed. The fact that pro-inflammatory S100 proteins contribute to *corpora amylacea* formation elevates their role as potential cancer risk factors. There is a growing body of evidence indicating that inflammation is a crucial prerequisite in prostate pathogenesis, as it is found to be associated with 40–90% of benign prostatic hyperplasia and with 20% of all human cancers [91]. Prostate cancer is the most common non-cutaneous malignant neoplasm in men in Western countries, affecting several million men in the Western world, and its incidence is rising rapidly with population ageing. Therefore, cancer risk assessment is of critical significance in its preventing strategies.

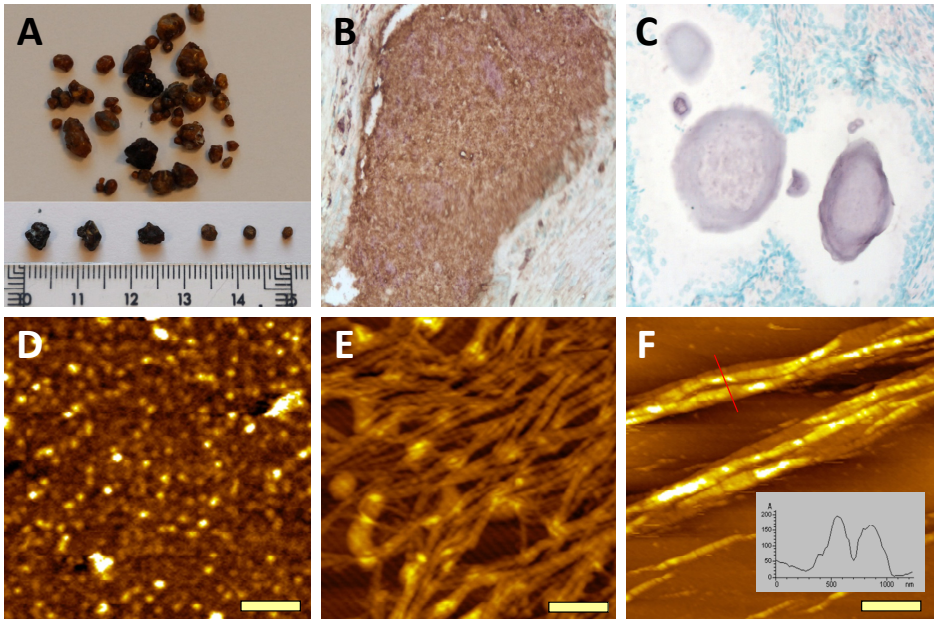


Figure 7.4 - Amyloid formation by S100A8/A9 proteins in the ageing prostate. (A) *Corpora amylacea* deposits extracted as result of prostatectomy (ruler is shown in centimeters). (B) Co-immunostaining of *corpora amylacea* with anti-S100A8 (shown in purple) and anti-S100A9 antibodies (shown in brown). (C) Immunostaining of *corpora amylacea* by antibodies towards amyloid fibrils (shown in purple). AFM images of (D) *ex vivo* amyloid oligomers; (E) ex vivo amyloid fibrillar network and (F) amyloid fibrils produced *in vitro* at pH 7.4, 37 °C with agitation. The fibril height analysis corresponds to the cross-section marked as a red line. Scale bars represent 250 nm.

By using mass-spectrometry, gel electrophoresis and Western blot analyses, we have found that proinflammatory S100A8/A9 proteins are persistently present in all specimens obtained as a result of prostatectomy in prostate cancer patients [15]. Immunohistochemical analysis of *corpora amylacea* revealed that they are stained positively with both anti-S100A8 and anti-S100A9 antibodies (Figure 2B). Positive foci of S100A8 and S100A9, including glandular epithelial cells and tissue macrophages, were observed in the tissues adjacent to *corpora amylacea* inclusions, indicating that the latter infiltrate inflamed glands and ultimately lead to raising local concentrations of S100A8/A9. Proteinaceous compounds constitutes up to 30-40 % of *corpora amylacea* deposits as revealed by X-ray photoelectron spectroscopy and FTIR, while the rest corresponds to inorganic components consisting of hydroxylapatite ($Ca_5(PO_4)_3OH$) and whitlockite ($Ca_2(PO_4)_3$), containing high concentration of Zn^{2+} ions. The calcification of protein deposits leads effectively to their further stabilisation in the protease-rich prostate fluid. The mineral content of *corpora amylacea* is rather uniform in all seven studied patients, indicating that calcification can be a regulated process. A recently reported function of S100A9 is associated with promoting calcification [92], suggesting that dystrophic calcification of *corpora amylacea* deposits could be influenced by the activities of S100A8/A9.

Remarkably, all *corpora amylacea* specimens were also stained with anti-amyloid fibril antibodies [93] (Figure 7.4C) and Congo red dye, used as a marker for the presence of the amyloid form of proteins, demonstrating that the amyloid material constitutes a significant mass of these specimens. Indeed, atomic force and transmission electron microscopy analyses revealed a variety of highly heterogeneous aggregates in the *corpora amylacea* extracts (Figure 7.4D-E), ranging from oligomeric species to extensive networks of mature fibrils, which is typical for the amyloid assemblies [94], as well as larger scale supramolecular assemblies, reaching

a few microns in length. Similar amyloid forms of S100A8/A9 were produced *in vitro*, providing further insight into their amyloidogenic properties. The S100A8/A9 complexes, extracted from granulocytes and produced recombinantly from *Escherichia coli*, were each incubated under the native conditions of pH 7.4 and 37 °C with agitation and at pH 2.0 and 57 °C without agitation. Under both conditions the proteins were assembled into heterogeneous fibrillar species. At pH 7.4, species resembling *ex vivo* oligomers and short protofilaments were formed after 2 weeks and thick bundles of fibrils with heights of 15–20 nm and a few microns in length constituted the major population of fibrillar aggregates after 8 weeks of incubation (Figure 7.4F). In the S100A8/A9 samples incubated at pH 2.0, oligomeric species and protofilaments also emerged in 2 weeks and flexible fibrils with height of ca. 4–5 nm and microns in length together with straight and rigid fibrillar structures a few hundred nanometers in length after 4 weeks, all closely resembling the ex vivo species.

It is important to note that Ca^{2+} and Zn^{2+} play a critical role in promoting amyloid assembly of S100A8/A9 proteins. As *ex vivo corpora amylacea* deposits are calcified and contain zinc salts, these ions can play a critical role in S100A8/A9 amyloid formation in vivo. Indeed, after 2 weeks of incubation the S100A8/A9 amyloid protofilaments of ca. 2 nm height were assembled in the presence of 10 mM $ZnCl_2$ and in a suspension of $Ca_3(PO_4)_2$ [15], but not when EDTA was added in solution. These species were converted into the fibrillar assemblies after 4 weeks of incubation, and again no filamentous structures were developed in the presence of EDTA.

The bundles of amyloid fibrils of S100A8/A9 proteins, formed both in vivo and *in vitro* (Figure 7.4F), are amongst the largest reported amyloid supramolecular species. The lateral association and thickening of the fibrils is likely to be a contributing factor to their stability in the prostate gland. It has been suggested that the various functions of the S100A8/A9 hetero- and

homo-oligomers may be regulated by their differential protease sensitivity [20]. The hetero-oligomeric complexes of S100A8/A9 are characterized by significant stability and protease resistance comparable to that of prions. In the protease rich environment of the prostate gland, and especially at sites of inflammation, where proteases are present at even higher levels, protease resistance of the S100A8/A9 proteins could favour their accumulation and conversion into amyloid structures. If so, the amyloid structures formed by the S100A8/A9 can be at the extreme end of the scale of resistance to proteolysis.

As prostatic fluid is very rich in protein content, small quantities of other proteins were also found in the *corpora amylacea* inclusions, presumably being trapped in the aggregating and growing deposits. Among them, the finding of *Escherichia coli* DNA and *Escherichia coli* proteins indicates that *corpora amylacea* formation may be associated with bacterial infection, causing consequently inflammation in surrounding tissues during the course of *corpora amylacea* establishment and growth. The identification of the highly amyloidogenic bacterial co-chaperonin GroES can be related not only to the fact that bacterial infection is a contributory factor to inflammation, but also suggests the potential role of bacterial infection in the initiating of the amyloid depositions via seeding [95].

As a result, a self-perpetuating cycle can be triggered in the ageing prostate, leading ultimately to amyloid growth. The increasing concentration of aggregation-prone proteins in the sites of inflammation would favor their amyloid assembly and deposition, as amyloid formation is a concentration-dependent process. This can be further promoted by the presence of calcium and zinc salts abundant in *corpora amylacea* and S100A8/A9 in turn can themselves regulate their own calcification. In the course of *corpora amylacea* growth, neighboring acini are obstructed, exacerbating inflammation and enhancing the risk of neoplastic transformation. Thus, the

direct involvement of pro-inflammatory S100A8/A9 proteins in *corpora amylacea* biogenesis emphasizes their role in the age-dependent prostate remodeling and accompanied ailments.

7.7. References

1. Moore BW (1965) A soluble protein characteristic of the nervous system. *Biochem Biophys Res Commun* **19**, 739-744
2. Engelkamp D, Schäfer BW, Mattei MG, Erne P & Heizmann CW (1993) Six S100 genes are clustered on human chromosome 1q21: identification of two genes coding for the two previously unreported calcium-binding proteins S100D and S100E. *Proc Natl Acad Sci U S A* **90**, 6547-6551
3. Schäfer BW, Wicki R, Engelkamp D, Mattei MG & Heizmann CW (1995) Isolation of a YAC clone covering a cluster of nine S100 genes on human chromosome 1q21: rationale for a new nomenclature of the S100 calcium-binding protein family. *Genomics* **25**, 638-643
4. Salama I, Malone PS, Mihaimeed F & Jones JL (2008) A review of the S100 proteins in cancer. *Eur J Surg Oncol* **34**, 357-364
5. Boom A, Pochet R, Authelet M, Pradier L, Borghgraef P, Van Leuven F, Heizmann CW & Brion JP (2004) Astrocytic calcium/zinc binding protein S100A6 over expression in Alzheimer's disease and in PS1/APP transgenic mice models. *Biochim Biophys Acta* **1742**, 161-168
6. Mrak RE & Griffinbc WS (2001) The role of activated astrocytes and of the neurotrophic cytokine S100B in the pathogenesis of Alzheimer's disease. *Neurobiol Aging* **22**, 915-922
7. Skelton NJ, Kordel J, Akke M, Forsen S & Chazin WJ (1994) Signal transduction versus buffering activity in Ca(2+)-binding proteins. *Nat Struct Biol* **1**, 239-245
8. Botelho HM, Koch M, Fritz G & Gomes CM (2009) Metal ions modulate the folding and stability of the tumor suppressor protein S100A2. *Febs J* **276**, 1776-1786
9. Heizmann CW & Cox JA (1998) New perspectives on S100 proteins: a multi-functional Ca(2+)-, Zn(2+)- and Cu(2+)-binding protein family. *Biometals* **11**, 383-397
10. Bjork P, Bjork A, Vogl T, Stenstrom M, Liberg D, Olsson A, Roth J, Ivars F & Leanderson T (2009) Identification of human S100A9 as a novel target for treatment of autoimmune disease via binding to quinoline-3-carboxamides. *PLoS Biol* **7**, e97
11. Leclerc E, Fritz G, Vetter SW & Heizmann CW (2009) Binding of S100 proteins to RAGE: an update. *Biochim Biophys Acta* **1793**, 993-1007
12. Ostendorp T, Leclerc E, Galichet A, Koch M, Demling N, Weigle B, Heizmann CW, Kroneck PM & Fritz G (2007) Structural and functional insights into RAGE activation by multimeric S100B. *Embo J* **26**, 3868-3878
13. Donato R (2007) RAGE: a single receptor for several ligands and different cellular responses: the case of certain S100 proteins. *Curr Mol Med* **7**, 711-724
14. Vogl T, Tenbrock K, Ludwig S, Leukert N, Ehrhardt C, van Zoelen MA, Nacken W, Foell D, van der Poll T, Sorg C, et al. (2007) Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock. *Nat Med* **13**, 1042-1049
15. Yanamandra K, Alexeyev O, Zamotin V, Srivastava V, Shchukarev A, Brorsson AC, Tartaglia GG, Vogl T, Kayed R, Wingsle G, et al. (2009) Amyloid formation by the pro-inflammatory S100A8/A9 proteins in the ageing prostate. *PLoS One* **4**, e5562
16. Marlatt NM, Boys BL, Konermann L & Shaw GS (2009) Formation of monomeric S100B and S100A11 proteins at low ionic strength. *Biochemistry* **48**, 1954-1963
17. Santamaria-Kisiel L, Rintala-Dempsey AC & Shaw GS (2006) Calcium-dependent and -independent interactions of the S100 protein family. *The Biochemical journal* **396**, 201-214
18. Korndörfer IP, Brueckner F & Skerra A (2007) The crystal structure of the human (S100A8/S100A9)2 heterotetramer, calprotectin, illustrates how conformational changes of interacting alpha-helices can determine specific association of two EF-hand proteins. *J Mol Biol* **370**, 887-898
19. Hunter MJ & Chazin WJ (1998) High level expression and dimer characterization of the S100 EF-hand proteins, migration inhibitory factor-related proteins 8 and 14. *J Biol Chem* **273**, 12427-12435

20. Nacken W & Kerkhoff C (2007) The hetero-oligomeric complex of the S100A8/S100A9 protein is extremely protease resistant. *FEBS Lett* **581**, 5127-5130
21. Vogl T, Leukert N, Barczyk K, Strupat K & Roth J (2006) Biophysical characterization of S100A8 and S100A9 in the absence and presence of bivalent cations. *Biochim Biophys Acta* **1763**, 1298-1306
22. Baudier J, Mandel P & Gerard D (1983) Bovine brain S100 proteins: separation and characterization of a new S100 protein species. *J Neurochem* **40**, 145-152
23. Yang Q, O'Hanlon D, Heizmann CW & Marks A (1999) Demonstration of heterodimer formation between S100B and S100A6 in the yeast two-hybrid system and human melanoma. *Exp Cell Res* **246**, 501-509
24. Deloulme JC, Assard N, Mbele GO, Mangin C, Kuwano R & Baudier J (2000) S100A6 and S100A11 are specific targets of the calcium- and zinc-binding S100B protein in vivo. *J Biol Chem* **275**, 35302-35310
25. Wang G, Rudland PS, White MR & Barraclough R (2000) Interaction in vivo and in vitro of the metastasis-inducing S100 protein, S100A4 (p9Ka) with S100A1. *J Biol Chem* **275**, 11141-11146
26. Wang G, Zhang S, Fernig DG, Spiller D, Martin-Fernandez M, Zhang H, Ding Y, Rao Z, Rudland PS & Barraclough R (2004) Heterodimeric interaction and interfaces of S100A1 and S100P. *The Biochemical journal* **382**, 375-383
27. Lehmann R, Melle C, Escher N & von Eggeling F (2005) Detection and identification of protein interactions of S100 proteins by ProteinChip technology. *J Proteome Res* **4**, 1717-1721
28. Moroz OV, Antson AA, Dodson EJ, Burrell HJ, Grist SJ, Lloyd RM, Maitland NJ, Dodson GG, Wilson KS, Lukanidin E, et al. (2002) The structure of S100A12 in a hexameric form and its proposed role in receptor signalling. *Acta Crystallogr D Biol Crystallogr* **58**, 407-413
29. Vogl T, Roth J, Sorg C, Hillenkamp F & Strupat K (1999) Calcium-induced noncovalently linked tetramers of MRP8 and MRP14 detected by ultraviolet matrix-assisted laser desorption/ionization mass spectrometry. *J Am Soc Mass Spectrom* **10**, 1124-1130
30. Novitskaya V, Grigorian M, Kriajevska M, Tarabykina S, Bronstein I, Berezin V, Bock E & Lukanidin E (2000) Oligomeric forms of the metastasis-related Mts1 (S100A4) protein stimulate neuronal differentiation in cultures of rat hippocampal neurons. *J Biol Chem* **275**, 41278-41286
31. Koch M, Bhattacharya S, Kehl T, Gimona M, Vasak M, Chazin W, Heizmann CW, Kroneck PM & Fritz G (2007) Implications on zinc binding to S100A2. *Biochim Biophys Acta* **1773**, 457-470
32. Rezvanpour A, Phillips JM & Shaw GS (2009) Design of high-affinity S100-target hybrid proteins. *Protein Sci* **18**, 2528-2536
33. Fritz G & Heizmann CW (2004) 3D structures of the calcium and zinc binding S100 proteins. In *Handbook of metalloproteins* (Messerschmidt A, Huber R, Poulos T & Wieghardt K, eds). John Wiley & Sons.
34. Capozzi F, Casadei F & Luchinat C (2006) EF-hand protein dynamics and evolution of calcium signal transduction: an NMR view. *J Biol Inorg Chem* **11**, 949-962
35. Biekofsky RR, Turjanski AG, Estrin DA, Feeney J & Pastore A (2004) Ab initio study of NMR 15N chemical shift differences induced by Ca2+ binding to EF-hand proteins. *Biochemistry* **43**, 6554-6564
36. Fritz G, Mittl PR, Vasak M, Grutter MG & Heizmann CW (2002) The crystal structure of metal-free human EF-hand protein S100A3 at 1.7-A resolution. *J Biol Chem* **277**, 33092-33098
37. Otterbein LR, Kordowska J, Witte-Hoffmann C, Wang CL & Dominguez R (2002) Crystal structures of S100A6 in the Ca(2+)-free and Ca(2+)-bound states: the calcium sensor mechanism of S100 proteins revealed at atomic resolution. *Structure* **10**, 557-567
38. Koch M, Diez J & Fritz G (2008) Crystal structure of Ca^{2+} -free S100A2 at 1.6 A resolution. *J Mol Biol* **378**, 933-942
39. Andersson M, Malmendal A, Linse S, Ivarsson I, Forsen S & Svensson LA (1997) Structural basis for the negative allostery between Ca(2+)- and Mg(2+)-binding in the intracellular Ca(2+)-receptor calbindin D9k. *Protein Sci* **6**, 1139-1147
40. Gribenko AV & Makhatadze GI (1998) Oligomerization and divalent ion binding properties of the S100P protein: a Ca2+/Mg2+-switch model. *J Mol Biol* **283**, 679-694
41. Matsuda S (1988) Ca2+-binding of S-100 protein in the presence of K+ and Mg2+. *J Biochem* **104**, 989-990
42. Pedrocchi M, Schafer BW, Mueller H, Eppenberger U & Heizmann CW (1994) Expression of Ca(2+)-binding proteins of the S100 family in malignant human breast-cancer cell lines and biopsy samples. *Int J Cancer* **57**, 684-690

43. Ridinger K, Schafer BW, Durussel I, Cox JA & Heizmann CW (2000) S100A13. Biochemical characterization and subcellular localization in different cell lines. *J Biol Chem* **275**, 8686-8694
44. Sturchler E, Cox JA, Durussel I, Weibel M & Heizmann CW (2006) S100A16, a novel calcium-binding protein of the EF-hand superfamily. *J Biol Chem* **281**, 38905-38917
45. Moroz OV, Blagova EV, Wilkinson AJ, Wilson KS & Bronstein IB (2009) The crystal structures of human S100A12 in apo form and in complex with zinc: new insights into S100A12 oligomerisation. *J Mol Biol* **391**, 536-551
46. Brodersen DE, Nyborg J & Kjeldgaard M (1999) Zinc-binding site of an S100 protein revealed. Two crystal structures of Ca2+-bound human psoriasin (S100A7) in the Zn2+-loaded and Zn2+-free states. *Biochemistry* **38**, 1695-1704
47. Charpentier TH, Wilder PT, Liriano MA, Varney KM, Pozharski E, MacKerell AD, Jr., Coop A, Toth EA & Weber DJ (2008) Divalent metal ion complexes of S100B in the absence and presence of pentamidine. *J Mol Biol* **382**, 56-73
48. Schäfer BW, Fritschy JM, Murmann P, Troxler H, Durussel I, Heizmann CW & Cox JA (2000) Brain S100A5 is a novel calcium-, zinc-, and copper ion-binding protein of the EF-hand superfamily. *J Biol Chem* **275**, 30623-30630
49. Moroz OV, Antson AA, Grist SJ, Maitland NJ, Dodson GG, Wilson KS, Lukanidin E & Bronstein IB (2003) Structure of the human S100A12-copper complex: implications for host-parasite defence. *Acta Crystallogr D Biol Crystallogr* **59**, 859-867
50. Landriscina M, Bagala C, Mandinova A, Soldi R, Micucci I, Bellum S, Prudovsky I & Maciag T (2001) Copper induces the assembly of a multiprotein aggregate implicated in the release of fibroblast growth factor 1 in response to stress. *J Biol Chem* **276**, 25549-25557
51. Nishikawa T, Lee IS, Shiraishi N, Ishikawa T, Ohta Y & Nishikimi M (1997) Identification of S100b protein as copper-binding protein and its suppression of copper-induced cell damage. *J Biol Chem* **272**, 23037-23041
52. Sivaraja V, Kumar TK, Rajalingam D, Graziani I, Prudovsky I & Yu C (2006) Copper binding affinity of S100A13, a key component of the FGF-1 nonclassical copper-dependent release complex. *Biophys J* **91**, 1832-1843
53. Glaser R, Harder J, Lange H, Bartels J, Christophers E & Schroder JM (2005) Antimicrobial psoriasin (S100A7) protects human skin from Escherichia coli infection. *Nat Immunol* **6**, 57-64
54. Baudier J, Glasser N & Gerard D (1986) Ions binding to S100 proteins. I. Calcium- and zinc-binding properties of bovine brain S100 alpha alpha, S100a (alpha beta), and S100b (beta beta) protein: Zn2+ regulates Ca2+ binding on S100b protein. *J Biol Chem* **261**, 8192-8203
55. Dell'Angelica EC, Schleicher CH & Santome JA (1994) Primary structure and binding properties of calgranulin C, a novel S100-like calcium-binding protein from pig granulocytes. *J Biol Chem* **269**, 28929-28936
56. Kerkhoff C, Vogl T, Nacken W, Sopalla C & Sorg C (1999) Zinc binding reverses the calcium-induced arachidonic acid-binding capacity of the S100A8/A9 protein complex. *FEBS Lett* **460**, 134-138
57. Moroz OV, Burkitt W, Wittkowski H, He W, Ianoul A, Novitskaya V, Xie J, Polyakova O, Lednev IK, Shekhtman A, et al. (2009) Both Ca2+ and Zn2+ are essential for S100A12 protein oligomerization and function. *BMC Biochem* **10**, 11
58. Baudier J & Cole RD (1988) Interactions between the microtubule-associated tau proteins and S100b regulate tau phosphorylation by the Ca2+/calmodulin-dependent protein kinase II. *J Biol Chem* **263**, 5876-5883
59. Yu WH & Fraser PE (2001) S100beta interaction with tau is promoted by zinc and inhibited by hyperphosphorylation in Alzheimer's disease. *J Neurosci* **21**, 2240-2246
60. Mbele GO, Deloulme JC, Gentil BJ, Delphin C, Ferro M, Garin J, Takahashi M & Baudier J (2002) The zinc- and calcium-binding S100B interacts and co-localizes with IQGAP1 during dynamic rearrangement of cell membranes. *J Biol Chem* **277**, 49998-50007
61. Gentil BJ, Delphin C, Mbele GO, Deloulme JC, Ferro M, Garin J & Baudier J (2001) The giant protein AHNAK is a specific target for the calcium- and zinc-binding S100B protein: potential implications for Ca2+ homeostasis regulation by S100B. *J Biol Chem* **276**, 23253-23261
62. Barber KR, McClintock KA, Jamieson GA, Jr., Dimlich RV & Shaw GS (1999) Specificity and Zn2+ enhancement of the S100B binding epitope TRTK-12. *J Biol Chem* **274**, 1502-1508
63. Ionescu JG, Novotny J, Stejskal V, Latsch A, Blaurock-Busch E & Eisenmann-Klein M (2006) Increased levels of transition metals in breast cancer tissue. *Neuro Endocrinol Lett* **27 Suppl 1**, 36-39

64. Li M, Zhang Y, Liu Z, Bharadwaj U, Wang H, Wang X, Zhang S, Liuzzi JP, Chang SM, Cousins RJ, et al. (2007) Aberrant expression of zinc transporter ZIP4 (SLC39A4) significantly contributes to human pancreatic cancer pathogenesis and progression. *Proc Natl Acad Sci U S A* **104**, 18636-18641
65. Xie J, Burz DS, He W, Bronstein IB, Lednev I & Shekhtman A (2007) Hexameric calgranulin C (S100A12) binds to the receptor for advanced glycated end products (RAGE) using symmetric hydrophobic target-binding patches. *J Biol Chem* **282**, 4218-4231
66. Kiryushko D, Novitskaya V, Soroka V, Klingelhofer J, Lukanidin E, Berezin V & Bock E (2006) Molecular mechanisms of Ca(2+) signaling in neurons induced by the S100A4 protein. *Molecular and cellular biology* **26**, 3625-3638
67. Mossberg A, Mok KH, Morozova-Roche LA & Svanborg C (2010) HAMLET-type complexes: structure and function. *FEBS J*, mini-review series
68. Zimmer DB & Van Eldik LJ (1986) Identification of a molecular target for the calcium-modulated protein S100. Fructose-1,6-bisphosphate aldolase. *J Biol Chem* **261**, 11424-11428
69. Brozzi F, Arcuri C, Giambanco I & Donato R (2009) S100B Protein Regulates Astrocyte Shape and Migration via Interaction with Src Kinase: IMPLICATIONS FOR ASTROCYTE DEVELOPMENT, ACTIVATION, AND TUMOR GROWTH. *J Biol Chem* **284**, 8797-8811
70. Millward TA, Heizmann CW, Schafer BW & Hemmings BA (1998) Calcium regulation of Ndr protein kinase mediated by S100 calcium-binding proteins. *Embo J* **17**, 5913-5922
71. Garbuglia M, Verzini M, Rustandi RR, Osterloh D, Weber DJ, Gerke V & Donato R (1999) Role of the C-terminal extension in the interaction of S100A1 with GFAP, tubulin, the S100A1- and S100B-inhibitory peptide, TRTK-12, and a peptide derived from p53, and the S100A1 inhibitory effect on GFAP polymerization. *Biochem Biophys Res Commun* **254**, 36-41
72. Donato R (1987) Quantitative analysis of the interaction between S-100 proteins and brain tubulin. *Cell Calcium* **8**, 283-297
73. Kilby PM, Van Eldik LJ & Roberts GC (1997) Identification of the binding site on S100B protein for the actin capping protein CapZ. *Protein Sci* **6**, 2494-2503
74. Wright NT, Cannon BR, Wilder PT, Morgan MT, Varney KM, Zimmer DB & Weber DJ (2009) Solution structure of S100A1 bound to the CapZ peptide (TRTK12). *J Mol Biol* **386**, 1265-1277
75. Austermann J, Nazmi AR, Heil A, Fritz G, Kolinski M, Filipek S & Gerke V (2009) Generation and characterization of a novel, permanently active S100P mutant. *Biochim Biophys Acta* **1793**, 1078-1085
76. Koltzscher M, Neumann C, Konig S & Gerke V (2003) Ca2+-dependent binding and activation of dormant ezrin by dimeric S100P. *Mol Biol Cell* **14**, 2372-2384
77. Baudier J, Delphin C, Grunwald D, Khochbin S & Lawrence JJ (1992) Characterization of the tumor suppressor protein p53 as a protein kinase C substrate and a S100b-binding protein. *Proc Natl Acad Sci U S A* **89**, 11627-11631
78. Mueller A, Schafer BW, Ferrari S, Weibel M, Makek M, Hochli M & Heizmann CW (2005) The calcium-binding protein S100A2 interacts with p53 and modulates its transcriptional activity. *J Biol Chem* **280**, 29186-29193
79. Grigorian M, Andresen S, Tulchinsky E, Kriajevska M, Carlberg C, Kruse C, Cohn M, Ambartsumian N, Christensen A, Selivanova G, et al. (2001) Tumor suppressor p53 protein is a new target for the metastasis-associated Mts1/S100A4 protein: functional consequences of their interaction. *J Biol Chem* **276**, 22699-22708
80. Ehrchen JM, Sunderkotter C, Foell D, Vogl T & Roth J (2009) The endogenous Toll-like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection, autoimmunity, and cancer. *J Leukoc Biol* **86**, 557-566
81. Foell D, Wittkowski H, Vogl T & Roth J (2007) S100 proteins expressed in phagocytes: a novel group of damage-associated molecular pattern molecules. *J Leukoc Biol* **81**, 28-37
82. Donato R, Sorci G, Riuzzi F, Arcuri C, Bianchi R, Brozzi F, Tubaro C & Giambanco I (2009) S100B's double life: intracellular regulator and extracellular signal. *Biochim Biophys Acta* **1793**, 1008-1022
83. Businaro R, Leone S, Fabrizi C, Sorci G, Donato R, Lauro GM & Fumagalli L (2006) S100B protects LAN-5 neuroblastoma cells against Abeta amyloid-induced neurotoxicity via RAGE engagement at low doses but increases Abeta amyloid neurotoxicity at high doses. *J Neurosci Res* **83**, 897-906
84. Turovskaya O, Foell D, Sinha P, Vogl T, Newlin R, Nayak J, Nguyen M, Olsson A, Nawroth PP, Bierhaus A, et al. (2008) RAGE, carboxylated glycans and S100A8/A9 play essential roles in colitis-associated carcinogenesis. *Carcinogenesis* **29**, 2035-2043

85. Ghavami S, Kerkhoff C, Chazin WJ, Kadkhoda K, Xiao W, Zuse A, Hashemi M, Eshraghi M, Schulze-Osthoff K, Klonisch T, et al. (2008) S100A8/9 induces cell death via a novel, RAGE-independent pathway that involves selective release of Smac/DIABLO and Omi/HtrA2. *Biochim Biophys Acta* **1783**, 297-311
86. Ghavami S, Eshragi M, Ande SR, Chazin WJ, Klonisch T, Halayko AJ, McNeill KD, Hashemi M, Kerkhoff C & Los M (2010) S100A8/A9 induces autophagy and apoptosis via ROS-mediated cross-talk between mitochondria and lysosomes that involves BNIP3. *Cell Res* **20**, 314-331
87. Viemann D, Barczyk K, Vogl T, Fischer U, Sunderkotter C, Schulze-Osthoff K & Roth J (2007) MRP8/MRP14 impairs endothelial integrity and induces a caspase-dependent and -independent cell death program. *Blood* **109**, 2453-2460
88. Loser K, Vogl T, Voskort M, Lueken A, Kupas V, Nacken W, Klenner L, Kuhn A, Foell D, Sorokin L, et al. (2010) The Toll-like receptor 4 ligands Mrp8 and Mrp14 are crucial in the development of autoreactive CD8+ T cells. *Nat med*, in press
89. Hiratsuka S, Watanabe A, Sakurai Y, Akashi-Takamura S, Ishibashi S, Miyake K, Shibuya M, Akira S, Aburatani H & Maru Y (2008) The S100A8-serum amyloid A3-TLR4 paracrine cascade establishes a pre-metastatic phase. *Nat Cell Biol* **10**, 1349-1355
90. Cross PA, Bartley CJ & McClure J (1992) Amyloid in prostatic corpora amylacea. *J Clin Pathol* **45**, 894-897
91. De Marzo AM, Platz EA, Sutcliffe S, Xu J, Gronberg H, Drake CG, Nakai Y, Isaacs WB & Nelson WG (2007) Inflammation in prostate carcinogenesis. *Nature reviews* **7**, 256-269
92. McCormick MM, Rahimi F, Bobryshev YV, Gaus K, Zreiqat H, Cai H, Lord RS & Geczy CL (2005) S100A8 and S100A9 in human arterial wall. Implications for atherogenesis. *J Biol Chem* **280**, 41521-41529
93. Kayed R, Head E, Sarsoza F, Saing T, Cotman CW, Necula M, Margol L, Wu J, Breydo L, Thompson JL, et al. (2007) Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol Neurodegener* **2**, 18
94. Pedersen JS, Andersen CB & Otzen DE (2010) Institutional nonconformism: the many levels of glucagon polymorphism. *FEBS J*, mini-review series
95. Otzen D & Nielsen PH (2008) We find them here, we find them there: functional bacterial amyloid. *Cell Mol Life Sci* **65**, 910-927

8. Metal ions modulate the folding and stability of the tumor suppressor protein S100A2

This chapter was published in

Botelho HM, Koch M, Fritz G & Gomes CM (2009) Metal ions modulate the folding and stability of the tumor suppressor protein S100A2. *FEBS J* 276, 1776-1786

Proteins were purified and supplied by Michael Koch and Günter Fritz (Konstanz University, DE).

8.1. Summary

The EF-hand protein S100A2 is a cell cycle regulator involved in tumorigenesis, acting through regulation of the p53 activation state. Metal ion free S100A2 is homodimeric and contains two Ca^{2+} binding sites and two Zn^{2+} binding sites per subunit, whereby the Zn^{2+} ion binding to one of the sites is coordinated by residues from two homodimers. The effect of selective binding of these metal ions was investigated using site specific mutants which lack one or both zinc sites. Circular dichroism analysis of secondary structure changes upon metallation have shown that Zn^{2+} binding is associated with a decrease in the secondary structure content, whereas Ca^{2+} has the opposite effect in two of the three S100A2 mutants studied. The energy of unfolding (ΔG_U) of the apo wild type S100A2 was determined to be 89.9 kJ/mol, and the apparent midpoint transition temperature (T_m^{app}) 58.4°C. In addition, a detailed study of the urea and thermal unfolding of the S100A2 mutants in different metallation states (apo, Zn^{2+} and Ca^{2+}) has been performed. Thermal denaturation experiments showed that Zn^{2+} acts as a destabilizer and Ca^{2+} as a stabilizer of the protein conformation. This suggests a synergistic effect between metal binding, protein stability and S100A2 biological activity, according to which Ca^{2+} activates and stabilizes the protein, the opposite being observed upon Zn^{2+} binding.

8.2. Introduction

S100A2 is a member of the S100 protein family, the largest subgroup within the superfamily of Ca^{2+}-binding EF-hand proteins. Human S100A2 is a 22 kDa homodimer, expressed mainly in the kidney, liver, heart and skeletal muscle [1]. Notably, the cellular localization of S100A2 is restricted to the nucleus [2-3]. S100A2 was found to be a tumor suppressor protein [4], which is down-regulated by promoter hypermethylation in breast and prostate cancers [5-6]. Its tumor suppressor activity is directly linked to p53, which is

activated by binding of S100A2, in a Ca^{2+}-dependent manner [7] ($K_d(Ca^{2+})$ ~ 100 μM). Each S100A2 protomer is composed of two tandem Ca^{2+}-binding helix-loop-helix EF-hands [8], of which the N-terminal one has a consensus sequence that is specific to S100 proteins (Figure 8.1). Like in other cases, the binding of Ca^{2+} to S100 proteins is known to induce structural changes: helix III rotates approximately 90º, exposing an inter-helical hydrophobic protein interaction site [9-11]. Zn^{2+} ions bind in two surface sites [12]. Site 1 has higher affinity and is composed of Cys21 and very likely His17, Gln22 and a solvent molecule. Site 2 comprises Cys2 and involves two S100A2 dimers. Both Ca^{2+} and Zn^{2+} are able to bind simultaneously to S100A2 as two Ca^{2+} binding events are detected when titrating the Zn^{2+} saturated protein [12]. Within the S100 family, Zn^{2+} has unique roles in S100A2 whose molecular basis remains to be established: *i.* Zn^{2+} binding is not common to family members and S100A2 exhibits the second highest Zn^{2+} affinity (K_d = 25 nM, close to S100A3, with K_d = 4 nM [13]), making S100A2 a more sensitive sensor for Zn^{2+}, rather than for Ca^{2+} ; *ii*. Zn^{2+} binding to the low affinity Cys2 site triggers dimerization of dimers, which is exclusive to S100A2; *iii.* physiologically relevant Zn^{2+} concentrations decrease the Ca^{2+} affinity, upon binding to the same Cys2 site [12]. The latter regulatory mechanism suggests that Zn^{2+}-loaded S100A2 ($K_{d,1}(Ca^{2+})$ = 832 μM $K_{d,2}(Ca^{2+})$ = 17 mM [12]) cannot activate p53, as physiological free Ca^{2+} concentrations do not exceed 100-300 μM [14-16]. This indicates that down-regulation of S100A2 takes place also at the

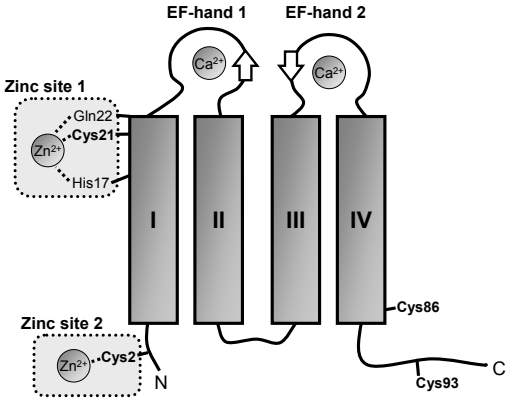


Figure 8.1 - S100A2 subunit topology, including the location of cysteines and other Zn^{2+}-coordinating residues [8, 12]. Boxes represent α-helices and arrows represent β-strands.

post-translational level.

In order to further explore the interplay between Zn^{2+} and Ca^{2+} binding to S100A2 and to address how metallation affects the protein conformation and stability we have investigated the effects of metal ions on the wild type protein and on mutants lacking one or both Zn^{2+}-binding sites. A detailed knowledge on how Zn^{2+} ions modulate the conformation and stability of S100A2 by will contribute to a better understanding of the regulation of protein function by metal ions, in particular as a putative Zn^{2+} sensor.

8.3. Materials and methods

8.3.1. Proteins

Wild type human S100A2 and mutants C2S and ΔCys (C2S-C21S-C86S-C93S) were expressed in *E. coli* and purified to homogeneity, as described elsewhere [12]. 2 mM Tris/HCl pH 7.0 was used throughout. All solutions were prepared in Chelex (Sigma, Steinheim, Germany) treated water and buffers were oxygen-free. Noteworthy, previous studies have determined that the cysteine to serine substitutions do not compromise the overall fold [12, 17].

8.3.2. Preparation of apo and metal loaded mutants

The proteins containing cysteines were reduced prior to all experiments as described elsewhere [12] and quantified spectrophotometrically ($\varepsilon^{275,\,wt}$ = 3050 $M^{-1}.cm^{-1}$; $\varepsilon^{280,\,C2S}$ = 3105 $M^{-1}.cm^{-1}$ and $\varepsilon^{280,\,\Delta Cys}$ = 2980 $M^{-1}.cm^{-1}$). Zn^{2+} was added as 1 or 2 molar equivalents to the S100A2 monomer in order to fill only the high affinity or both sites. Ca^{2+} was added as 10 molar equivalents. Metal chloride salts were used (Fluka, Steinheim, Germany). For CD and fluorescence measurements in the presence of metals, the protein samples were equilibrated 1 h at 4°C, after the addition of the metal.

8.3.3. Circular dichroism spectroscopy

CD measurements were recorded in a Jasco J-815 spectropolarimeter equipped with a Peltier-controlled thermostated cell support. Thermal denaturation experiments were carried out increasing the temperature from 25 to 95°C at a heating rate of 1°C/min. Changes of the CD signal at 222 nm were plotted as a function of temperature and the T_m^{app} was determined from fitting to single or the sum of two sigmoidal curves. Protein concentration was 0.1 mg/ml. Thermal denaturation was irreversible. However, no kinetically controlled steps affected protein unfolding as the T_m^{app} was independent of the heating rate for S100A2-ΔCys (not shown), like observed in a system undergoing reversible unfolding. Therefore, the determined T_m^{app} are those of a pseudo-equilibrium and are suited for comparative purposes between the mutants studied.

The thermal denaturation kinetics (25-65°C temperature jumps) of the single Zn^{2+} site mutant S100A2-C2S was investigated in the apo and Zn^{2+}-loaded (1:1) state, following the decay of the CD signal at 225 nm. This mutant preserves Zn^{2+} site 1, present in all S100A2 mutants, and does not tetramerize because it lacks site 2. Protein concentration was 0.2 mg/ml.

8.3.4. Fluorescence spectroscopy

Intrinsic tyrosine fluorescence measurements were performed on a Varian Cary Eclipse instrument. Temperature was kept at 25°C by a Peltier-controlled cell support. Emission spectra upon 275 nm excitation were recorded using 10 nm excitation and emission slits.

8.3.5. ATR FT-IR spectroscopy

ATR FT-IR measurements were performed in a Bruker IFS 66/S spectrometer equipped with a nitrogen-cooled MCT detector using the thermostatized Harrick BioATRcell II. Spectra were acquired at a 4 cm^{-1} resolution. Difference spectra were calculated after vector normalization of

the absorbance in the amide I – amide II region. Different metallated forms of S100A2 mutants were prepared by in situ dialysis using the manufacturer's accessory. Apo protein samples (~10 mg/ml) were dialysed at 20°C against 5 mM Tris/HCl pH 7 23 mM NaCl. 10 molar equivalents of Ca^{2+} were added in the same buffer with NaCl concentration adjusted to equalize the ionic strength.

Thermal denaturation experiments consisted of increasing the cell temperature from 20 to 94°C. The heating rate was discontinuous (average ~1.3°C/min) as a result of stepwise spectrum acquisition (every 2°C, acquisition time: 1 min). Denaturation curves were obtained by plotting spectra 2^{nd} derivative values at local maxima or minima as a function of temperature.

8.3.6. Chemical denaturation

Protein unfolding was studied by monitoring the variation of CD at 222 nm or fluorescence intensity at 305 nm, at 25°C, as a function of urea concentration. Fresh urea (Riedel-de Haën, Seelze, Germany) solutions were used for every assay and the rigorous concentration was determined using refractive index measurements [18]. Apo or metallated protein samples (0.1 mg/ml) were incubated 2 h at room temperature for complete chemical denaturation. The influence of protein concentration on the C_m was assessed in the 0.08-0.25 mg/ml range. Denaturation was reversible for all cases as urea dilution of the completely denatured protein yielded protein with native state spectra.

8.3.7. Dynamic light scattering

The molecular diameter of S100A2 mutants in different metallation conditions (0.1 mg/ml) was assessed using a Malvern Instruments Zetasizer Nano ZS instrument equipped with a 633 nm laser. Temperature was kept at 25°C using a Peltier-controlled thermostatized cell support. Before each

measurement, samples were filtered through a 0.22 μM membrane. For each time measurement, the backscattered light (173°) from fourteen 10 s accumulations was averaged. Results were analyzed with Malvern Instruments DTS software using a multimodal fit with quadratic weighting and 0.01 regulariser. Size results are from the Mie theory derived volume distribution of sizes. When available, error bars are the standard deviations from at least 3 replicate measurements.

8.4. Results and discussion

8.4.1. Structural changes upon Ca^{2+} and Zn^{2+} binding

In order to investigate the effect of Ca^{2+} and Zn^{2+} ions on the structure of S100A2, two previously characterized mutants [12] were studied together with the wild type protein. Cysteine residues, which are part of the two S100A2 Zn^{2+} sites (Figure 8.1), have been replaced by serines in mutants C2S and ΔCys (all 4 cysteines in each subunit are replaced by serines). Each mutant has therefore a different number of available Zn^{2+} sites: whereas S100A2-wt has two sites, S100A2-C2S only preserves one high affinity site and S100A2-ΔCys is devoid of specific Zn^{2+} sites. These mutations do not affect the Ca^{2+} affinity [12], thus allowing analyzing the Zn^{2+} role upon binding to the available sites.

These S100A2 mutants were investigated in the apo and in holo forms corresponding to different metallated states at 25°C using far-UV circular dichroism (CD) (Figure 8.2). The CD spectra of all protein preparations are typical of α-helix proteins, with local minima at 208 and 222 nm and local maxima at 195 and 215 nm, in agreement with the ΔCys-S100A2 crystal structure [8] and other structural data [12, 17, 19-21]. This observation also corroborates the previous observation that the cysteine replacements do not affect the overall protein fold [12].

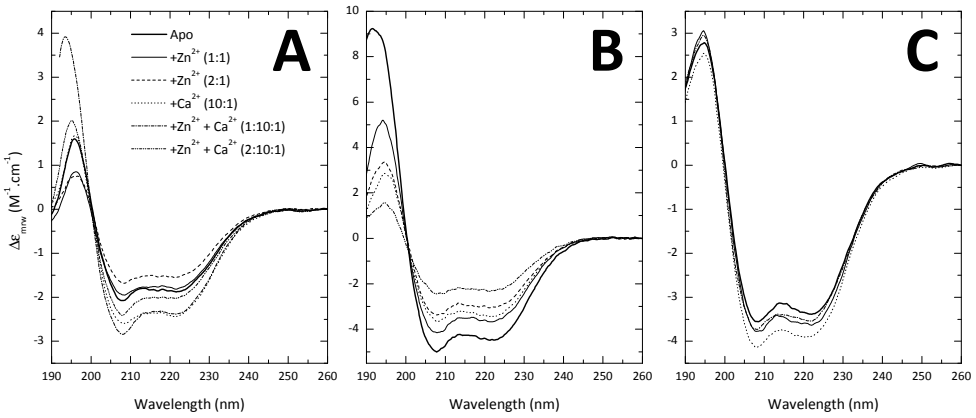


Figure 8.2 - CD spectra of S100A2 wt (A), C2S (B) and ΔCys (C) in several metal load conditions: apo (thick solid line), 1 Eq Zn^{2+} (dot), 2 Eq Zn^{2+} (dash), 10 Eq Ca^{2+} (thin solid line), 1 Eq Zn^{2+} plus 10 Eq Ca^{2+} (dash dot) and 2 Eq Zn^{2+} plus 10 Eq Ca^{2+} (dash dot dot).

Binding of Zn^{2+} to S100A2-wt and S100A2-ΔCys does not elicit significant secondary structure changes (Figure 8.2AC). In the latter case, this is justified by the absence of Zn^{2+} sites, although the far-UV CD spectrum is sensitive to unspecific Zn^{2+} binding to this mutant (data not shown). On the other hand, Zn^{2+} binding to the S100A2-C2S mutant produced a concentration-dependent decrease of secondary structure (Figure 8.2B) upon addition of 1 and 2 Zn^{2+} equivalents, respectively. Binding of Ca^{2+} to S100A2-wt and S100A2-ΔCys results in an increase in the α-helical content (Figure 8.2AC). An opposite effect is observed in the S100A2-C2S mutant (Figure 8.2B). The highest increase in secondary structure occurs when both Ca^{2+} and Zn^{2+} are added to the wild type protein (Figure 8.2A).

In order to investigate the possibility that the observed variations in secondary structure resulting from metallation with Ca^{2+} and Zn^{2+} are due to a change in the oligomeric state of the proteins or aggregation, we carried out dynamic light scattering (DLS) studies. We have detected average molecular diameters around 5-5.5 nm, irrespectively of mutation or metal load (Figure 8.3), consistent with the structure of apo S100A2 [8]. The

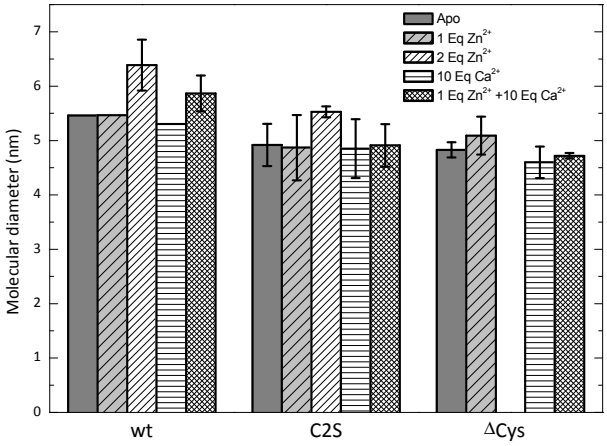


Figure 8.3 - Molecular diameter of the apo and metallated S100A2 variants as assessed by DLS. When available, results are average ± standard deviation of at least 3 replicate measurements. Sizes are derived from the volume-averaged distribution of sizes. There is no measurement of S100A2-ΔCys +2 Eq Zn^{2+} because the protein aggregates.

diameter of S100A2-wt +2 Eq Zn^{2+} is slightly higher (6.4 nm), reflecting partial tetramerization.

8.4.2. Chemical stability of holo and apo S100A2 proteins

The conformational stability of S100A2-wt and mutants, in the apo and distinct metallated states was investigated by performing urea denaturation experiments. For all proteins, the far-UV CD spectra obtained at increasing urea concentrations denoted a transition from α-helix to random conformations, apparently via intermediate β-sheet structures (Figure 8.4A).

To extract thermodynamic information from protein denaturation curves the unfolding mechanism has to be known. For single domain dimeric proteins like S100A2, this process may be hypothesized to comprise two steps: the dissociation of the native dimer into folded monomers which in turn undergoes denaturation. However, the chemical denaturation of S100A2 could be rationalized using a simple two state unfolding mechanism, where the unfolding of the folded dimer (F_2) yields denatured monomer (U) directly:

$$F_2 \rightleftharpoons 2U \qquad (1)$$

This mechanism is supported by several criteria: i. No intermediate species were detected in none of the denaturation curves (Figure 8.4); ii. The C_m of apo wt and S100A2-ΔCys increased with protein concentration (not shown), and; iii. Denaturation curves of the latter mutant obtained by CD and intrinsic tyrosine fluorescence were superimposable (not shown) [22]. Accordingly, the mechanism in equation 1 was used to derive the thermodynamic parameters, using the formalism established by Grant et al. [23] (Figure 8.4B-D, Table 8.1, Table 8.2).

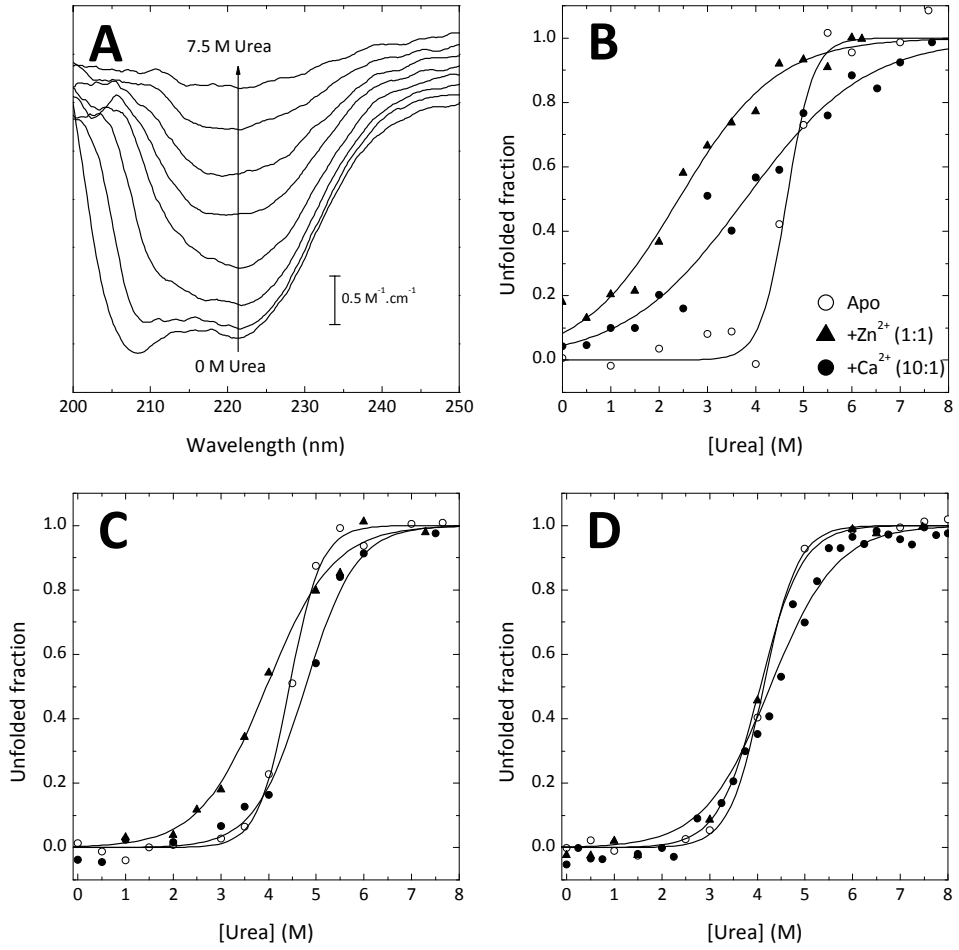


Figure 8.4 - Circular dichroism-monitored urea chemical denaturation curves of S100A2. Representative spectra of Ca^{2+}-loaded S100A2-wt at increasing urea concentration (0 - 7.5 M), as indicated by the arrow (A). Displacement: 0.1 $M^{-1}.cm^{-1}$. Denaturation curves of wt (B), C2S (C) and S100A2-ΔCys (D) in several metal load conditions: apo (○), 1 Eq Zn^{2+} (▲) and 10 Eq Ca^{2+} (●).

Table 8.1 - Thermodynamic stability parameters for the apo S100A2 variants

	$[Urea]_{1/2}$ (M)	ΔG_U (kJ/mol)	m (kJ.mol^{-1} M^{-1})	$\Delta\Delta G_U$[a] (kJ/mol)
Wt	4.7	89.9	13.0	-
C2S	4.5	78.9	11.2	-2.3
ΔCys	4.2	71.8	10.3	-5.8

[a]$\Delta\Delta G_U = \Delta[Urea]_{1/2} \times m_{average}$ [18]

Table 8.2 - Thermodynamic stability parameters for the S100A2 variants in the apo, Zn^{2+} (1:1) and Ca^{2+} (10:1) metallated states.

	S100A2-wt				S100A2-C2S				S100A2- ΔCys			
	$[Urea]_{½}$ (M)	ΔG_U (kJ/mol)	m (kJ.mol^{-1}.M^{-1})	$\Delta\Delta G_U$[a] (kJ/mol)	$[Urea]_{½}$ (M)	ΔG_U (kJ/mol)	m (kJ.mol^{-1}.M^{-1})	$\Delta\Delta G_U$[a] (kJ/mol)	$[Urea]_{½}$ (M)	ΔG_U (kJ/mol)	m (kJ.mol^{-1}.M^{-1})	$\Delta\Delta G_U$[a] (kJmol)
Apo	4.7	89.9	13.0	-	4.5	78.9	11.2	-	4.2	71.8	10.3	-
Zn^{2+}	2.5	35.1	3.5	-14.2	4.0	50.1	5.3	-3.9	4.1	68.2	9.7	-0.8
Ca^{2+}	3.9	40.3	2.9	-5.2	4.8	62.4	7.0	+2.4	4.3	51.8	5.3	+0.8

[a]$\Delta\Delta G_U = \Delta[Urea]_{1/2} \times m_{average}$ [18]

A two state unfolding mechanism has also been reported for human S100B [24] and porcine S100A12 [25]. All stability parameters extracted from denaturation curves (Figure 8.4B-D, Table 8.1, Table 8.2) were found to be within the typical range for small dimeric proteins [26] and, in particular, in accordance to thermodynamic data reported on human S100B [24] and porcine S100A12 [25].

The ΔG_U of the apo S100A2-wt was 89.9 kJ/mol and S100A2-C2S and ΔCys were destabilized by -2.3 and -5.8 kJ/mol in respect to the wild type (Table 8.1). The data suggest identical unfolding mechanisms for all three apo proteins as neither the transition cooperativity (m index) nor the shape of the far-UV CD spectra at different urea concentrations (not shown) were significantly affected. Thus, meaningful information on the thermodynamic stability of the Ca^{2+}- or Zn^{2+}- loaded mutants can be retrieved from analyzing

metallation effects within the background of the same mutation (Table 8.2). With the exception of Zn^{2+}-loaded S100A2-ΔCys, which is devoid of Zn^{2+} sites, the metallated states exhibit a decreased cooperativity of the unfolding transition. This suggests that in such cases the amount of surface area being exposed during urea unfolding is lower than in the apo state and/or that metal binding increases the subpopulations of native protein with slightly different conformations.

However, occupation of the metal sites by Zn^{2+} or Ca^{2+} ions has a distinct effect on the protein stability. Metallation of the high affinity Zn^{2+} site of S100A2-C2S has a destabilizing effect of -3.9 kJ/mol, while the same stoichiometric Zn^{2+} amount destabilizes the wild type protein by -14.2 kJ/mol. This large destabilization probably arises from residual binding of Zn^{2+} to the low affinity site, which is known to promote exposure of hydrophobic surfaces [12]. In contrast, binding of Ca^{2+} stabilizes the mutants by +0.8 or +2.4 kJ/mol, while destabilizing the wild type protein by -5.2 kJ/mol. Some point mutations are known to exert long range effects in S100 proteins because of their effect in hydrogen bond networks [27]. It is reasonable to hypothesize that the same applies to the S100A2 mutants under study. The lower unfolding cooperativity of the Ca^{2+}-loaded samples suggests a concurrent opening of the EF-hands, resulting in a decreased exposure of the surface area during unfolding.

8.4.3. Thermal stability of holo and apo S100A2 proteins

We have complemented the chemical denaturation study by performing analogous temperature-induced unfolding assays. For all proteins, increasing the temperature results in a progressive α-helix to random coil transition (Figure 8.5A). No notorious protein precipitation was observed, suggesting that other non-reversible modification may occur at high temperatures. This differed from the urea unfolding, precluding a detailed thermodynamic analysis, and is suggestive of distinct pathways for chemical and thermal

unfolding. Nevertheless, a comparison of the apparent midpoint transition temperature (T_m^{app}) obtained for the different metallated states is very informative in respect to the effect of each metal in the stability of each protein mutant (Table 8.3).

Apo S100A2-wt and S100A2-ΔCys have a very similar T_m^{app} of 58.4°C and 59.5°C, respectively, which are lower than the T_m^{app} value of 66.6°C of S100A2-C2S. The outlying behavior of apo S100A2-C2S may result from long

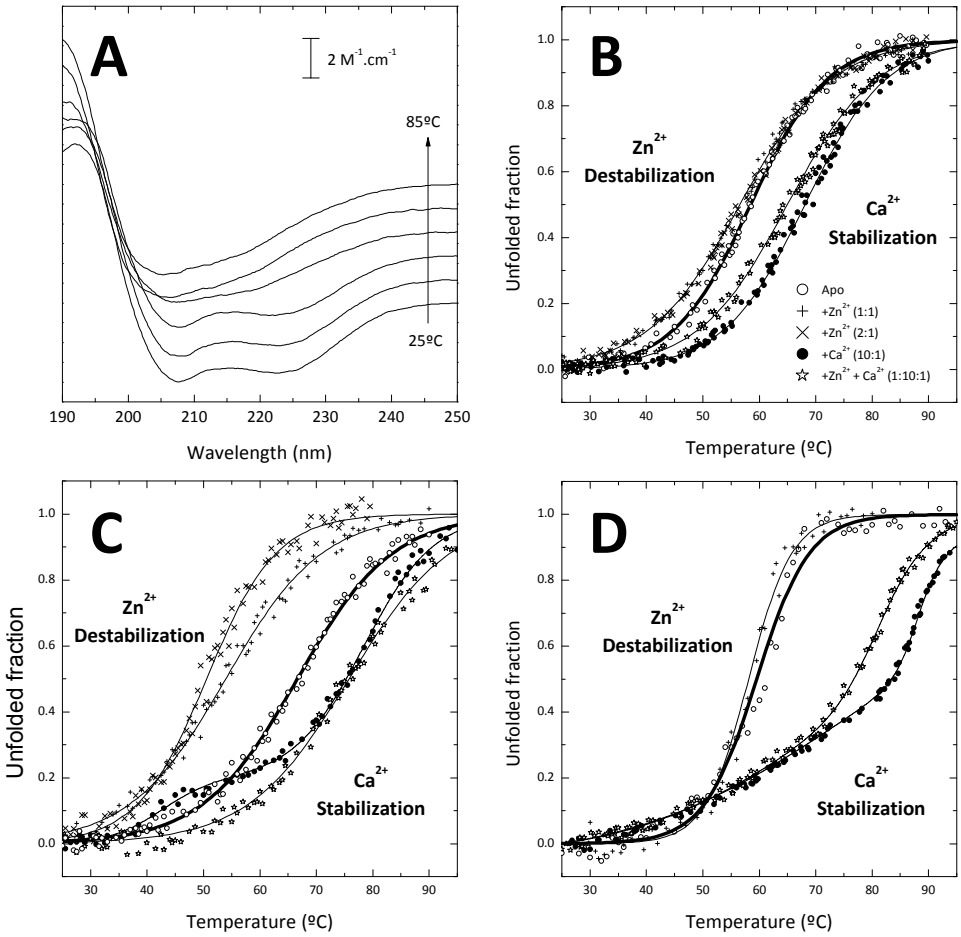


Figure 8.5 - Circular dichroism-monitored thermal denaturation of S100A2. Representative spectra of apo S100A2-wt at increasing temperatures (25-85°C), as indicated by the arrow (A). Displacement: 1.2 $M^{-1}.cm^{-1}$. The native and denatured form spectra are representative for all thermal denaturations. Thermal denaturation curves of wt (B), C2S (C) and S100A2-ΔCys (D) in several metal load conditions: apo (○), 1 Eq Zn^{2+} (+), 2 Eq Zn^{2+} (×), 10 Eq Ca^{2+} (●) and 1 Eq Zn^{2+} plus 10 Eq Ca^{2+} (✯).

Table 8.3 - Apparent T_m values determined from CD-monitored thermal denaturation curves of S100A2 variants. The aggregation of S100A2-ΔCys incubated with 2 Eq Zn^{2+} occurs during the temperature ramp.

	T_m^{app} (°C)					
	Apo	**+Zn^{2+} (1:1)**	**+Zn^{2+} (2:1)**	**+Ca^{2+} (10:1)**	**+Zn^{2+} +Ca^{2+} (1:10:1)**	**+Zn^{2+} +Ca^{2+} (2:10:1)**
wt	58.4	56.6	56.6	68.1	65.0	n. d.
C2S	66.6	53.9	50.6	~42 > 80	n. d.	76.8
ΔCys	59.5	58.1	Aggregates	~75 > 85	~65 > 81	Aggregates

range mutation effects [27], which are not observed in the other mutants. Considering these aspects, the relevant comparisons will relate with differences observed upon selective metallation, within the same mutant.

Interestingly, Ca^{2+} and Zn^{2+} metallation showed antagonistic effects in thermal stability (Figure 8.5B-D). Zn^{2+} ions had a destabilizing effect, which was concentration dependent in S100A2-C2S, in agreement with the observed decrease in secondary structure content. The destabilization arose from the metal-induced conformational change because no kinetic distortions affected the Zn^{2+}-induced conformational destabilization. Binding of Zn^{2+} to the unfolded state could have caused cause a shift in the equilibrium, but this effect is only significant at a large excess of Zn^{2+} [28], which is not the case in our experiments (a maximum of 1 or 2 Zn^{2+} equivalents were used). Also, the kinetics of thermal denaturation did not vary significantly between apo, and Zn^{2+}-loaded S100A2-C2S (see methods). The Ca^{2+}-loaded proteins exhibit an increased T_m^{app}, although the mutants have at least one unfolding intermediate in the denaturation curves. The increased stability of Ca^{2+}-loaded proteins likely resides in the electrostatic compensation at the negatively charged Ca^{2+}-binding sites.

The opposite effect of both metals on the thermal stability prompted us to study Ca^{2+} and Zn^{2+} ions in combination. In S100A2-wt, where all binding

sites are available, metal ion effects are dominated by the Ca^{2+} contribution. In fact, an intermediate stability in respect to the Zn^{2+} destabilization (ΔT_m^{app}= -1.8°C) and Ca^{2+} stabilization (ΔT_m^{app}= +9.7°C) is determined when Zn^{2+} and Ca^{2+} are combined (ΔT_m^{app}= +6.6°C) (Figure 8.5B).

It can be hypothesized that the two thermal transitions of Ca^{2+}-loaded S100A2-C2S correspond to unfolding of different structural regions, the transition at ~ 42°C (Figure 8.5C) corresponding to unfolding at the N-terminal EF-hand, as it is stabilized by Zn^{2+} binding at the adjacent site 1 (Figure 8.1). Such stabilization is not observed in S100A2-ΔCys (Figure 8.5D), which has no specific Zn^{2+} sites. In this case, unspecific Zn^{2+} binding is likely to result in destabilization without changing the shape of the denaturation curve.

Complementarily to the CD experiments, the thermal denaturation of S100A2-ΔCys was followed by FT-IR. The absorbance change at the amide I (1600-1700 cm^{-1}) and amide II (1500-1600 cm^{-1}) bands was used to probe the unfolding, monitoring secondary structure elements. As shown above, this mutant does not bind Zn^{2+}, so we have carried out a study of the apo and Ca^{2+}-loaded forms of ΔCys-S100A. In both conditions, denaturation consisted of transition from α-helical (~1650 and 1550 cm^{-1}) to random (1525 cm^{-1}) and β structures (~1622 cm^{-1}) (Figure 8.6AC). The latter vibration is associated with intermolecular β-sheets and aggregation. The formation of insoluble β-sheet containing aggregates is most certainly an important contributor to the irreversibility of the thermal denaturation. The denaturation curves of the above-mentioned structural elements are compatible with the CD results, and further corroborate a two-state unfolding process. All secondary structure elements of apo S100A2-ΔCys exhibit similar profiles, with T_m^{app} ranging from 67-71°C. The unfolding of the secondary structure elements of Ca^{2+}-loaded S100A2-ΔCys also occurs

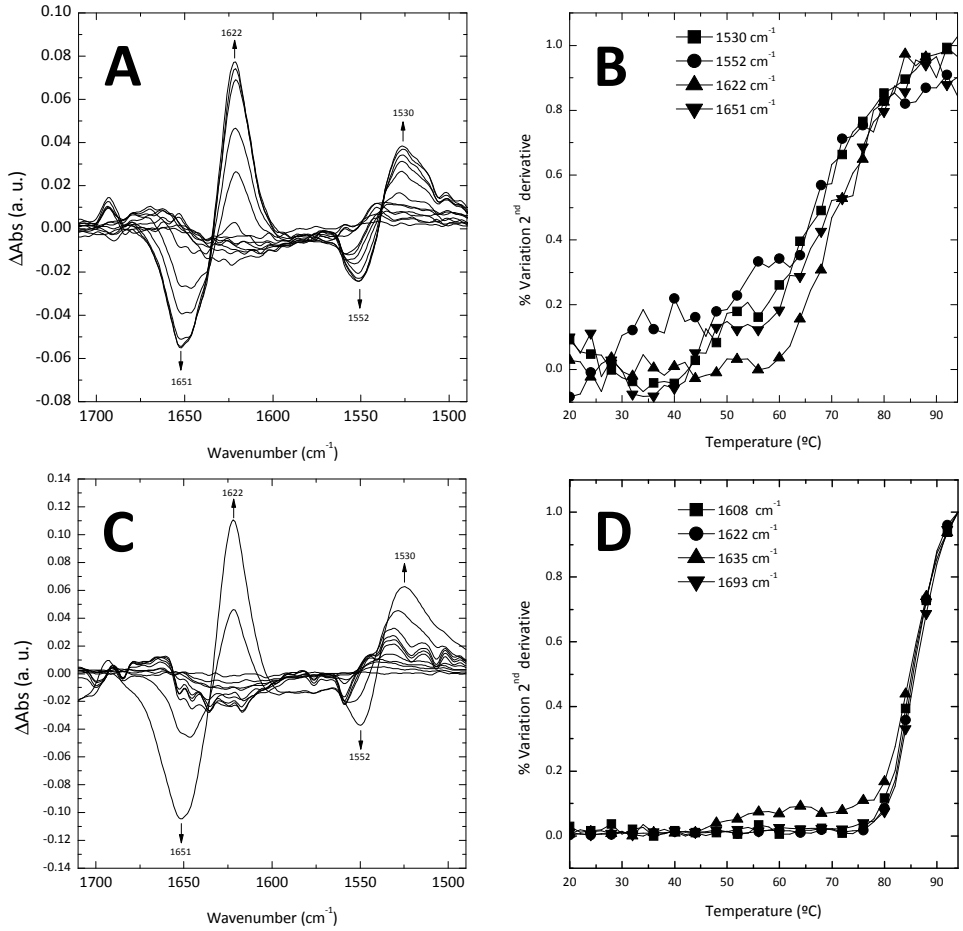


Figure 8.6 - ATR FT-IR-monitored thermal denaturation of S100A2-ΔCys in the apo (A,B) and Ca^{2+}-loaded forms (C,D). Representative difference spectra at increasing temperature (20-94°C), as indicated by the arrows (A and C). Thermal denaturation curves for the apo (B) and Ca^{2+}-loaded protein (D) are derived from the 2nd derivative trend with temperature.

simultaneously, and at a T_m^{app} > 80°C. Again, a very good agreement with the far-UV CD data is observed.

8.5. Conclusions

In this work we have characterized how the conformation and stability of S100A2 is influenced by specific metal ions, Zn^{2+} and Ca^{2+}. In particular, considering the unique role of Zn^{2+} in S100A2, we have dissected the contribution arising from Zn^{2+} binding using two mutants, which selectively

disrupt the low and the high affinity Zn^{2+}-binding sites, as models. We have observed that the S100A2 conformation is sensitive to the metallation state, although it retains essentially an α-helical structure. Chemical denaturation suggests that both Zn^{2+} and Ca^{2+}-associated conformational changes facilitate the accessibility of urea to the protein core, leading to destabilization. Thermal denaturation suggest that Zn^{2+} and Ca^{2+} regulate protein thermal stability antagonistically, being Zn^{2+} a destabilizer and Ca^{2+} a stabilizer. Similarly, Ca^{2+} stabilizes and Cu^{2+} destabilizes S100A13 towards thermal perturbation [29]. Other studies highlight distinct regulatory mechanisms of S100 proteins by metals ions. For example, Ca^{2+} was shown to stabilize human S100B towards denaturation by guanidinium hydrochloride [23] and porcine S100A12 was shown to be stabilized by Ca^{2+} and Zn^{2+} towards thermal denaturation [25]. The behavior of Zn^{2+}-Ca^{2+}-loaded S100A2 in the thermal unfolding experiments indicates that Ca^{2+} can at least partially revert the conformational destabilization triggered by Zn^{2+} binding to the high affinity site.

These effects of metal ions on S100A2 folding and stability contribute to a better understanding to the Ca^{2+} and Zn^{2+} dependent regulation of the protein. In the Ca^{2+}-loaded state S100A2 binds and activates p53 [7]. However, Zn^{2+} negatively regulates the affinity of S100A2 for Ca^{2+} binding [12], which might disable the Ca^{2+} signal, resulting in a blockage of p53 activation. The mechanism of how Zn^{2+} may decrease the Ca^{2+} affinity remained unresolved in our previous study [12]. The results of the present work reveal that the decrease in Ca^{2+} affinity through Zn^{2+} is presumably a result of the general destabilization of the protein. Further contributions might come from the exposure of a hydrophobic surface on Zn^{2+} binding [12], making additional exposure of the hydrophobic surface induced by Ca^{2+} less favourable.

Zn^{2+} binding to close homologues of S100A2, such as S100A3 [13] and S100A4 (G. Fritz and M. Koch, unpublished data), also occurs mainly via cysteine residues. It remains to be shown whether Zn^{2+} binding to S100A3 and S100A4 also results in a decrease in protein stability. In S100A3, Zn^{2+} binding causes the loss of approximately 40% of the a-helical structure [13], supporting destabilization of the protein. In contrast with S100A2, other S100 proteins, such as S100A12 and S100B, display an increased Ca^{2+} affinity on Zn^{2+} binding [30-31]. Future investigations might show whether, in these S100 proteins, Zn^{2+} increases the conformational stability, thereby facilitating the Ca^{2+} conformational change.

Together, the data presented here provide new insights into the mechanism of Zn^{2+}- and Ca^{2+}- dependent activation of S100 proteins. The antagonistic effect of Zn^{2+} and Ca^{2+} in the control of S100A2 stability provides a molecular rationale for the action of both metal ions. Our results allow the formulation of the following hypothesis: in tissues expressing S100A2, the Zn^{2+} imbalance which arises in some cancers may contribute to enhanced cell proliferation through destabilization of S100A2. This mechanism would impair the interaction with p53, and disrupt subsequent downstream cell cycle regulation. Indeed, Zn^{2+} transporters are upregulated in breast carcinoma and pancreatic tumours [32-33] leading to elevated Zn^{2+} levels [34-36], which may impair Ca^{2+} binding to S100A2 [12]. Current work in our laboratories will allow the testing of this hypothesis.

8.6. Acknowledgements

This work was supported by grants POCTI/QUI/45758 and PTDC/QUI/70101 (to CMG) from the Fundação para a Ciência e a Tecnologia (FCT/MCTES, Portugal), DAAD D/07/13610 PPP (to GF and MK). CMG and GF are recipients of a CRUP/DAAD collaborative grant A-15/08. HMB is a recipient of a Ph. D. fellowship (SFRH/BD/31126/2006) from Fundação para a Ciência e a Tecnologia (FCT/MCTES, Portugal).

8.7. References

1. Glenney JR, Jr., Kindy MS & Zokas L (1989) Isolation of a new member of the S100 protein family: amino acid sequence, tissue, and subcellular distribution. *J Cell Biol* **108**, 569-578
2. Mueller A, Bachi T, Höchli M, Schäfer BW & Heizmann CW (1999) Subcellular distribution of S100 proteins in tumor cells and their relocation in response to calcium activation. *Histochem Cell Biol* **111**, 453-459
3. Mandinova A, Atar D, Schäfer BW, Spiess M, Aebi U & Heizmann CW (1998) Distinct subcellular localization of calcium binding S100 proteins in human smooth muscle cells and their relocation in response to rises in intracellular calcium. *J Cell Sci* **111 (Pt 14)**, 2043-2054
4. Lee SW, Tomasetto C & Sager R (1991) Positive selection of candidate tumor-suppressor genes by subtractive hybridization. *Proc Natl Acad Sci U S A* **88**, 2825-2829
5. Pedrocchi M, Schäfer BW, Mueller H, Eppenberger U & Heizmann CW (1994) Expression of Ca(2+)-binding proteins of the S100 family in malignant human breast-cancer cell lines and biopsy samples. *Int J Cancer* **57**, 684-690
6. Rehman I, Cross SS, Catto JW, Leiblich A, Mukherjee A, Azzouzi AR, Leung HY & Hamdy FC (2005) Promoter hyper-methylation of calcium binding proteins S100A6 and S100A2 in human prostate cancer. *Prostate* **65**, 322-330
7. Mueller A, Schäfer BW, Ferrari S, Weibel M, Makek M, Höchli M & Heizmann CW (2005) The calcium-binding protein S100A2 interacts with p53 and modulates its transcriptional activity. *J Biol Chem* **280**, 29186-29193
8. Koch M, Diez J & Fritz G (2008) Crystal structure of Ca2+ -free S100A2 at 1.6-Å resolution. *J Mol Biol* **378**, 931-940
9. Fritz G & Heizmann CW (2004) 3D structures of the calcium and zinc binding S100 proteins. In *Handbook of metalloproteins* (Messerschmidt A, Huber R, Poulos T & Wieghardt K, eds). John Wiley & Sons.
10. Heizmann CW & Cox JA (1998) New perspectives on S100 proteins: a multi-functional Ca(2+)-, Zn(2+)- and Cu(2+)-binding protein family. *Biometals* **11**, 383-397
11. Lewit-Bentley A & Rety S (2000) EF-hand calcium-binding proteins. *Curr Opin Struct Biol* **10**, 637-643
12. Koch M, Bhattacharya S, Kehl T, Gimona M, Vasak M, Chazin W, Heizmann CW, Kroneck PM & Fritz G (2007) Implications on zinc binding to S100A2. *Biochim Biophys Acta* **1773**, 457-470
13. Fritz G, Mittl PR, Vasak M, Grutter MG & Heizmann CW (2002) The crystal structure of metal-free human EF-hand protein S100A3 at 1.7-A resolution. *J Biol Chem* **277**, 33092-33098
14. Alonso MT, Villalobos C, Chamero P, Alvarez J & Garcia-Sancho J (2006) Calcium microdomains in mitochondria and nucleus. *Cell Calcium* **40**, 513-525
15. Jones HE, Holland IB, Baker HL & Campbell AK (1999) Slow changes in cytosolic free Ca2+ in Escherichia coli highlight two putative influx mechanisms in response to changes in extracellular calcium. *Cell Calcium* **25**, 265-274
16. Knight MR, Campbell AK, Smith SM & Trewavas AJ (1991) Recombinant aequorin as a probe for cytosolic free Ca2+ in Escherichia coli. *FEBS Lett* **282**, 405-408
17. Koch M, Diez J & Fritz G (2006) Purification and crystallization of the human EF-hand tumour suppressor protein S100A2. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 1120-1123
18. Pace CN, Shirley BA & Thomson JA (1990) Measuring the conformational stability of a protein. In *Protein structure - a practical approach* (Creighton TE, ed, pp. 311-330. IRL Press, Oxford.
19. Franz C, Durussel I, Cox JA, Schafer BW & Heizmann CW (1998) Binding of Ca2+ and Zn2+ to human nuclear S100A2 and mutant proteins. *J Biol Chem* **273**, 18826-18834
20. Randazzo A, Acklin C, Schäfer BW, Heizmann CW & Chazin WJ (2001) Structural insight into human Zn(2+)-bound S100A2 from NMR and homology modeling. *Biochem Biophys Res Commun* **288**, 462-467
21. Stradal TB, Troxler H, Heizmann CW & Gimona M (2000) Mapping the zinc ligands of S100A2 by site-directed mutagenesis. *J Biol Chem* **275**, 13219-13227
22. Bosshard H, R., (2005) Concurrent Association and Folding of Small Oligomeric Proteins. In *Protein Folding Handbook* (Buchner J & Kiefhaber T, eds), pp. 965-997. Wiley-VCH Verlag GmbH, Weinheim.
23. Grant SK, Deckman IC, Culp JS, Minnich MD, Brooks IS, Hensley P, Debouck C & Meek TD (1992) Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases. *Biochemistry* **31**, 9491-9501

24. Ferguson PL & Shaw GS (2002) Role of the N-terminal helix I for dimerization and stability of the calcium-binding protein S100B. *Biochemistry* **41**, 3637-3646
25. Garcia AF, Garcia W, Nonato MC & Araujo AP (2008) Structural stability and reversible unfolding of recombinant porcine S100A12. *Biophys Chem* **134**, 246-253
26. Mei G, Di Venere A, Rosato N & Finazzi-Agro A (2005) The importance of being dimeric. *Febs J* **272**, 16-27
27. Nelson MR, Thulin E, Fagan PA, Forsen S & Chazin WJ (2002) The EF-hand domain: a globally cooperative structural unit. *Protein Sci* **11**, 198-205
28. Conejero-Lara F, Mateo PL, Aviles FX & Sanchez-Ruiz JM (1991) Effect of Zn2+ on the thermal denaturation of carboxypeptidase B. *Biochemistry* **30**, 2067-2072
29. Sivaraja V, Kumar TK, Rajalingam D, Graziani I, Prudovsky I & Yu C (2006) Copper binding affinity of S100A13, a key component of the FGF-1 nonclassical copper-dependent release complex. *Biophys J* **91**, 1832-1843
30. Baudier J, Glasser N & Gerard D (1986) Ions binding to S100 proteins. I. Calcium- and zinc-binding properties of bovine brain S100 alpha alpha, S100a (alpha beta), and S100b (beta beta) protein: Zn2+ regulates Ca2+ binding on S100b protein. *J Biol Chem* **261**, 8192-8203
31. Dell'Angelica EC, Schleicher CH & Santome JA (1994) Primary structure and binding properties of calgranulin C, a novel S100-like calcium-binding protein from pig granulocytes. *J Biol Chem* **269**, 28929-28936
32. Li M, Zhang Y, Liu Z, Bharadwaj U, Wang H, Wang X, Zhang S, Liuzzi JP, Chang SM, Cousins RJ, et al. (2007) Aberrant expression of zinc transporter ZIP4 (SLC39A4) significantly contributes to human pancreatic cancer pathogenesis and progression. *Proc Natl Acad Sci U S A* **104**, 18636-18641
33. Taylor KM, Morgan HE, Smart K, Zahari NM, Pumford S, Ellis IO, Robertson JF & Nicholson RI (2007) The emerging role of the LIV-1 subfamily of zinc transporters in breast cancer. *Mol Med* **13**, 396-406
34. Cui Y, Vogt S, Olson N, Glass AG & Rohan TE (2007) Levels of zinc, selenium, calcium, and iron in benign breast tissue and risk of subsequent breast cancer. *Cancer Epidemiol Biomarkers Prev* **16**, 1682-1685
35. Geraki K, Farquharson MJ & Bradley DA (2002) Concentrations of Fe, Cu and Zn in breast tissue: a synchrotron XRF study. *Phys Med Biol* **47**, 2327-2339
36. Ionescu JG, Novotny J, Stejskal V, Latsch A, Blaurock-Busch E & Eisenmann-Klein M (2006) Increased levels of transition metals in breast cancer tissue. *Neuro Endocrinol Lett* **27 Suppl 1**, 36-39

9. S100: A family of proteins with new amyloid-forming properties with potential implications in human disease

Part of this chapter was published in

Fritz G, Botelho HM, Morozova-Roche LA, Gomes CM (2010) Natural and amyloid self-assembly of S100 proteins: structural basis of functional diversity. *FEBS J*, 277(22) 4578-4590

Proteins were purified by Günter Fritz (Freiburg University, DE). AFM analysis was performed by Kiran Yanamandra and Ludmilla A. Morozova-Roche (Umeå University, SE)

9.1. Summary

S100 proteins constitute a protein family with structurally homologous elements exhibiting a high degree of conformational plasticity. They are signaling proteins coupling calcium binding to conformational changes and interaction with other proteins, sometimes requiring functional oligomerization. In this respect, the recently identified amyloid forming properties of S100A8/A9 raised the possibility of a new level of conformational regulation in these proteins. We have assessed the amyloidogenic potential of selected human S100 proteins using thioflavin T fluorescence enhancement and FT-IR spectroscopy and found that apo S100A2, S100A3, S100A4, S100A6 and S100B promptly form amyloid structures under acidic conditions. Atomic force microscopy imaging has revealed distinct amyloid structures for each protein: oligomeric (S100A6, S100B), protofibrilllar (S100A2) and fibrillar (S100A3). S100A12, which did not form ThT-reactive species, nevertheless formed aggregates. Here, we dissect the amyloidogenic processes of S100A6 and S100B using FT-IR, thioflavin T fluorescence and circular dichroism. These results are framed in respect to the phylogeny of S100 proteins and in the computed amyloidogenic potential. We propose that members of the S100 family are intrinsically amyloidogenic, a phenomenon with possible implications in human pathology.

9.2. Introduction

The S100 protein family is a remarkable example of conformational plasticity, wherein the binding of metal ions, most frequently calcium, triggers conformational transitions rendering the protein capable of binding other signaling polypeptides and fulfill distinct functions depending of the partner. Subtler effects occur when metal binding at specific sites influences binding properties elsewhere (*e. g.* positive calcium binding cooperativity or

the influence of zinc on calcium binding affinity). Within specific S100 proteins, dimerization may be promiscuous, with S100A1, S100A4, S100A6, S100A7, S100A8, S100A9, S100A10, S100A11, S100B and S100P accommodating heterodimerization partners. Moreover, the dimeric S100 building block may also form higher order assemblies, namely tetramers, hexamers and octamers. Recently, novel conformational properties of S100 proteins were uncovered as S100A8/A9 were implicated in amyloidogenic processes in the prostate [1]. Calcium and zinc-rich amyloid deposits named c*orpora amylacea* were shown to be mainly composed of S100A8 and S100A9 in the amyloid form. The co-deposition of bacterial proteins and macrophage recruitment suggested that S100 amyloidogenesis could be associated with inflammatory processes, associated with prostate cancer onset. The importance of metal ions in the amyloidogenic conversion was established as pathological-like protein fibrils could be assembled *in vitro* in the obligatory presence of calcium or zinc. Importantly, *corpora amylacea* exist in the normal brain and are mainly composed of S100 proteins, raising the possibility of S100 fibrillization being a general process with possible implication in other pathologies.

The S100 protein family is composed of more than 20 members in humans. Despite remarkable sequence and structural similarities [2], S100 proteins exhibit a high degree of functional specialization, including specific cell and tissue expression patterns. Due to their altered expression level changes, they are frequently used as clinical markers for cancer, although mis-regulation is also common in cardiomyopathy, inflammation and neurodegeneration [3]. In accordance, if the amyloidogenic properties of S100A8/A9 were shared by other family members, there would be putative implications in these diseases.

Fibrillization in amyloid structures has been suggested to be a general feature of all proteins [4]. The distinct amyloid-permissive conditions for

each protein are then responsible for the different amyloid propensities among different proteins. Structural studies have revealed a staggering similarity of fibrils built by different proteins at the microscopic as well as at the molecular level. These observations have allowed understanding the ubiquitous fibrillization accessibility as the outcome of being the backbone and not the amino acid side chains the responsible for the interactions polymerizing individual polypeptides in the fibril. Nevertheless, like all biophysical properties of proteins, the amyloidogenic propensity depends on the overall protein chemistry, set up by its sequence. No single physicochemical property can account for the amyloidogenic propensity of a protein [5]. Currently, the best assessments of amyloidogenic potential are based either on the physicochemical amino acid properties (*e. g.* TANGO [6]), the combination of amyloid-prone pattern recognition with physicochemical properties and homology modeling (*e. g.* WALTZ) or the correlation of aggregation rate changes upon mutation and polypeptide physicochemical properties (*e. g.* Zyggregator [7]). For the characterization of amyloid processes, numerous biochemical, biophysical and spectroscopic methodologies can be used to probe the varying properties of proteins during the amyloid assembly process.

Based on published data on the biology of S100 proteins and the recently described amyloidogenic properties of S100A8/A9 we have analyzed the amyloid formation propensity of the human S100 proteins, finding widespread amyloid-prone regions in several family members. From this initial result, we have screened several S100 family members with distinct biochemical and functional properties for their amyloidogenic potential at acidic pH. We have observed that several S100 proteins promptly form amyloid structures with very distinct morphologies. We report a structural characterization of the tested S100 amyloids. We discuss the putative

widespread amyloidogenic nature of the S100 family and its implications in disease.

9.3. Materials and methods

9.3.1. Chemicals and proteins

All reagents were of the highest grade commercially available. Thioflavin T was obtained from Sigma. Trace metals were removed from all solutions by solution treatment with a chelex resin. S100 proteins (S100A2, S100A3, S100A4, S100A6, S100A12 and S100B) were expressed in *E. coli* and purified to homogeneity as described previously [8]. Molar protein concentrations refer to the S100 subunit. Hen egg white lysozyme was obtained from Fluka.

9.3.2. Amyloid propensity assessment

Amyloid formation propensities of human S100 proteins at pH 7 were assessed using the Zyggregator (http://www-vendruscolo.ch.cam.ac.uk/zyggregator_test.php) [7, 9] and Waltz (http://waltz.vub.ac.be/) online servers. In Zyggregator, both the intrinsic amyloid aggregation propensity (Z_{agg}) and the oligomerization tendency (Z_{tox}) were retrieved. In Waltz, best overall performance settings were selected. For data comparing, S100 sequences were aligned in ClustalX [10] and GeneDoc [11].

9.3.3. Amyloid formation assays

S100 proteins were placed in 50 mM Glycine buffer pH 2.5 by repeated dilution and re-concentration. Sample temperature was kept at 4°C at all times before amyloid formation assays. Protein concentration was set at 3 mg/ml using Bradford's method [12]. Due to spontaneous and extensive aggregation in aerobic conditions, S100A3 (4 mg/ml) was previously reduced with 25 mM TCEP overnight. Any pre-existing aggregates were pelleted and removed by centrifuging at 12000 g for 10 minutes. Amyloid

formation was promoted by quiescent incubation at 57°C. The positive control for amyloid formation was hen egg white lysozyme, at 10 mg/ml in the same buffer and incubated at 57°C, a previously validated assay [13]. The structural characterization of S100B amyloid formation was carried out at 10mg/ml protein concentration.

9.3.4. Thioflavin T fluorescence

ThT binding assays were performed according to [13]. Briefly, 5 µl of protein sample were mixed with 65 µM ThT in 10 mM potassium phosphate (KPi), 150 mM NaCl, pH 7.0 at room temperature to a final volume of 1 ml. Samples were incubated under stirring for 2 min and ThT fluorescence at 482 nm was recorded with a Cary Varian Eclipse instrument with temperature set at 25°C using a peltier-thermostated cell support. Excitation wavelength was 440 nm, PMT voltage 600 V, and excitation and emission slits 10 nm.

9.3.5. Circular dichroism spectroscopy

The folding and stability of S100A6 (0.1 mg/ml) at pH 7 (50 mM KPi) and pH 2.5 (50 mM glycine) were compared through far UV CD spectroscopy. Measurements were recorded in a Jasco J-815 spectropolarimeter equipped with a Peltier-controlled thermostated cell support. Thermal denaturation experiments were carried out increasing the temperature from 20 to 95°C at a heating rate of 1°C/min. The normalized Changes of the CD signal at 222 nm was plotted as a function of temperature and the T_m^{app} was determined from a sigmoidal fit. Denaturation was irreversible. Spectra in the 180-260 nm wavelength range were acquired at 25°C before and after heating, as well as at 95°C. Spectra were deconvoluted with CDNN.

9.3.6. Atomic force microscopy

After reaching the plateau phase in amyloid formation kinetics, S100 samples were stored at 4°C and morphology was analyzed by AFM on a PICO PLUS 5500 microscope (Agilent) as previously described [14].

9.3.7. Dynamic light scattering

For monitoring molecular size variations during amyloid formation at 57°C, S100 proteins (3 mg/ml in 50 mM glycine pH 2.5) were filtered through a 0.22 μM membrane and analyzed with a Malvern Instruments Zetasizer Nano ZS instrument equipped with a 633 nm laser. Temperature was kept using a Peltier-controlled thermostatized cell support. Kinetic data was obtained from continuously repeated measurements. In each measurement, the backscattered light (173°) from fifteen 10 s accumulations was averaged. Amyloid formation was assessed from light scattering intensity and the average particle size obtained from the intensity distribution. Results were analyzed with Malvern Instruments DTS software using a multimodal fit with quadratic weighting and 0.01 regulariser.

9.3.8. ATR FT-IR spectroscopy

Secondary structural changes occurring during S100B amyloid formation were monitored by ATR FT-IR measurements performed in a Bruker IFS 66/S spectrometer equipped with a nitrogen-cooled MCT detector using the thermostatized Harrick BioATR II cell. 10 mg/ml S100B in 50 mM glycine pH 2.5 were incubated at 57°C while spectra were continuously acquired (1 min accumulation time, 12 mm aperture, 20 kHz scanner velocity). Spectral resolution was 4 cm^{-1}. Amyloid formation was assessed by the absorption ratio between the intermolecular β-sheet characteristic band at 1618 cm^{-1} and the amide I absorption maximum, occurring between 1642 and 1652 cm^{-1}. Secondary structure quantification was based in the spectral deconvolution of the amide I band with lorentzian curves centered at second derivative

maxima of the self-deconvoluted spectra. Assignment of each lorentzian to secondary structure features took into consideration typical absorption regions [15]. Curve integration yielded the secondary structure estimations.

9.4. Results

9.4.1. Amyloidogenic potential of S100 proteins

The amyloidogenic potential of a protein can be estimated at the amino acid residue level using different algorithms. We have used the WALTZ [16] and Zyggregator [7, 9] algorithms to calculate the intrinsic aggregation propensity of human S100 proteins. WALTZ computes the amyloid formation probability (as a percentage). Zyggregator can compute the fibrillization potential (Z_{agg} score) as well as the propensity to form toxic β-rich oligomers (Z_{tox} score). If the difference between the amyloidogenic potential of the query sequence and the one of a random sequence is expressed and the standard deviation, its value is the Zyggregator score. Accordingly, significant scores are the ones larger than the unity. In addition, Zyggregator scores can be averaged to yield an overall propensity. We have set up our calculations at pH 7.0 to assess the amyloidogenic at near physiological conditions (Figure 9.1).

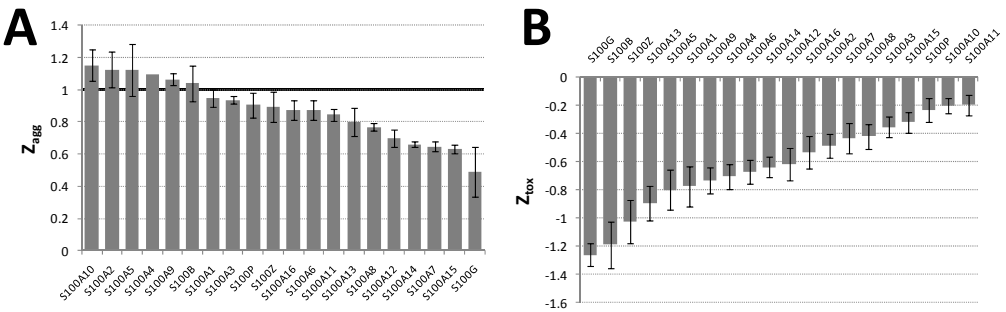


Figure 9.1 – Overall amyloid formation propensity of S100 proteins. Sequence-averaged scores, computed at pH 7.0. (A) Intrinsic aggregation propensity (Z_{agg}). (B) Oligomeric assembly formation propensity (Z_{tox}). Bars represent the score error.

Interestingly, for most S100 family members the fibrillization propensity (Z_{agg}) is comparable with the one of Aβ peptides, forming amyloid deposits in Alzheimer's disease: the overall aggregation scores for $A\beta_{1-40}$ and $A\beta_{1-42}$ at

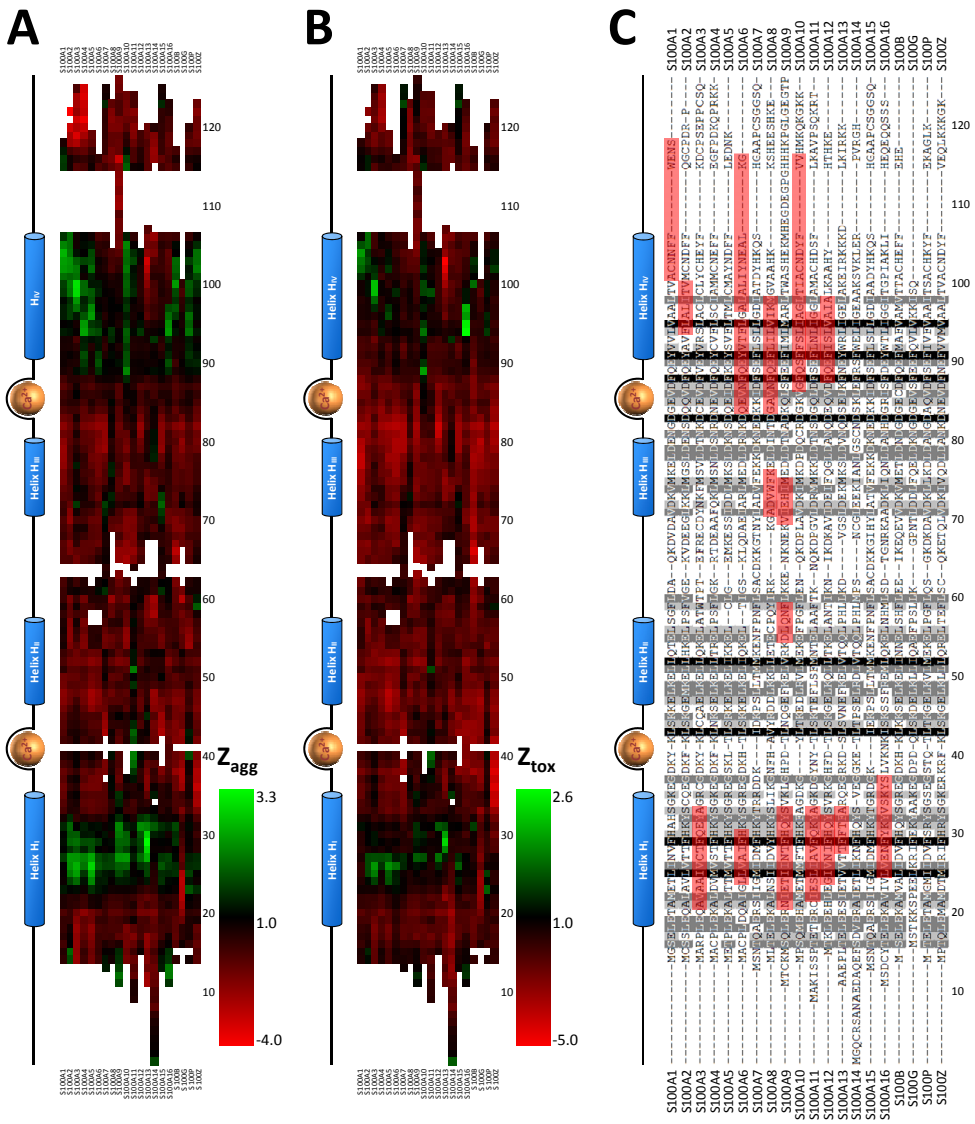


Figure 9.2 – Position-specific amyloidogenic potential of S100 proteins. Alignment of human S100 protein sequences, showing the intrinsic aggregation propensity (Z_{agg}) (A), the oligomeric assembly formation propensity (Z_{tox}) (B) or the WALTZ high score regions (C) at pH 7.0. Zyggregator results (A-B) are color coded: values above the significance threshold (>1) are in shades of green. WALTZ scores (C) either equal zero or are above 92% (red shaded sequences). Shades of grey highlight sequence consensus. Cartoons represent the canonical location of secondary structure elements in S100 proteins. Cylinders represent helices and spheres represent the position of EF-hand calcium binding loops.

pH 7.0 are 0.97 and 0.94, respectively [9]. These results thus evidence a rather high fibrillization propensity in the S100 family. The overall Z_{tox} scores are low. To determine which protein sequences are contributing to this amyloidogenic propensity, we determined the position-specific Zyggregator and WALTZ scores (Figure 9.2).

The position-specific Zyggregator scores (Figure 9.2AB) reveal that amyloidogenic regions are segregated at helices H_I and H_{IV}, which form the hydrophobic core of S100 proteins. Every family member displays a distinct location of amyloidogenic regions, not all of them being located at H_I and H_{IV} sites, but all family members display highly amyloidogenic regions. Interestingly, in contrast with the low overall values, almost all S100 proteins display regions with highly significant oligomerization propensities (Z_{tox} >> 1). Exceptions are S100A8 and S100G. The overall low Z_{tox} values are thus the consequence of the highly clustered oligomerization-prone regions in a mostly non-amyloidogenic context. The WALTZ score in S100 sequences was found to be either 0% or above 92%. Accordingly, we have defined WALTZ amyloidogenic and non-amyloidogenic sequences (red zones in Figure 9.2C), which are found in most S100 proteins: S100A1, S100A2, S100A3, S100A6, S100A8, S10A9, S100A10, S100A11, S100A12, S100A13, S100A16. More importantly, the Waltz amyloidogenic sequences also cluster around helix H_I and H_{IV}, in accordance with the Zyggregator analysis.

9.4.2. S100 proteins form amyloid-like species in vitro

Building on these initial observations and considering the fact that S100 proteins share a rather high chemical and structural identity, we have further addressed the hypothesis that amyloid formation could be a generalized property among the members of the S100 protein family. For this purpose, we have carried out a series of preliminary experiments in conditions identical to those assayed for S100A8/A9 (pH 2.5 and 57°C) to test if other S100 proteins (S100A2, S100A3, S100A4, S100A6, S100A12 and S100B)

would form thioflavin-T (ThT)-reactive amyloid species (Figure 9.3). These proteins are representative of the overall biological roles of S100 proteins: cell proliferation control /cancer (S100A2, S100A4, S100A6, S100B) [17], inflammation (S100A12, S100B) [18], cell motility (S100A4, S100B) [2], chemotaxis (S100A2, S100A12, S100B) [19-21] and RAGE activation (S100A2, S100A4, S100A6, S100A12, S100B) [22]. Lysozyme, a well studied amyloid-forming protein [13] was used as a positive control.

After prolonged incubation (≥50 h) all tested S100 proteins – with the exception of S100A12 – promptly formed ThT-binding amyloid structures that resulted in increase of fluorescence intensity of ThT, comparable to that observed upon dye interaction with the lysozyme amyloids used as a positive control. The kinetics of amyloidogenesis was similar in all amyloid-forming S100 proteins. Near endpoint fluorescence values were reached within the first hours of incubation. Three kinds of behavior are observed: the ThT negative S100A12; the proteins forming amyloid from a ThT negative state, putatively native-like (S100A4, S100A6, S100B); and proteins exhibiting a

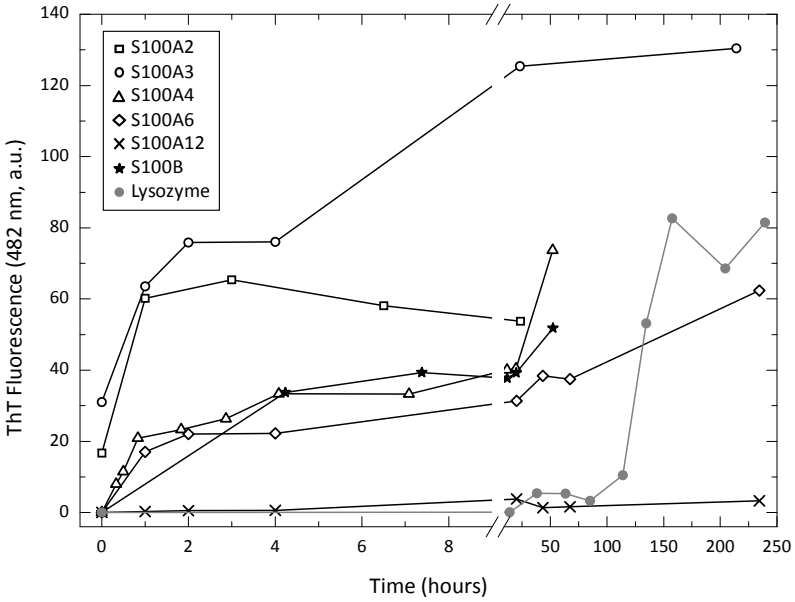


Figure 9.3 – Amyloid formation kinetics by S100 proteins.

ThT positive signal even before thermal treatment (S100A2, S100A3). The latter case suggests that S100A2 and S100A3 are highly amyloidogenic, being particularly susceptible to pH acidification which can induce amyloid formation by itself. No lag phase was detected in any of the amyloid-forming samples, an indication of the highly amyloidogenic nature of these proteins.

9.4.3. Morphology of S100 amyloids

The characterization of amyloid structures requires a morphological description which complements the ThT information on enhanced β-sheet content. After reaching the plateau phase in the amyloid formation assays (Figure 9.3), S100 samples were adhered to mica and analyzed by atomic force microscopy (AFM), which allowed the detection of amyloid and other precursor structures (fibrils, protofibrils and disordered aggregates) (Figure 9.4).

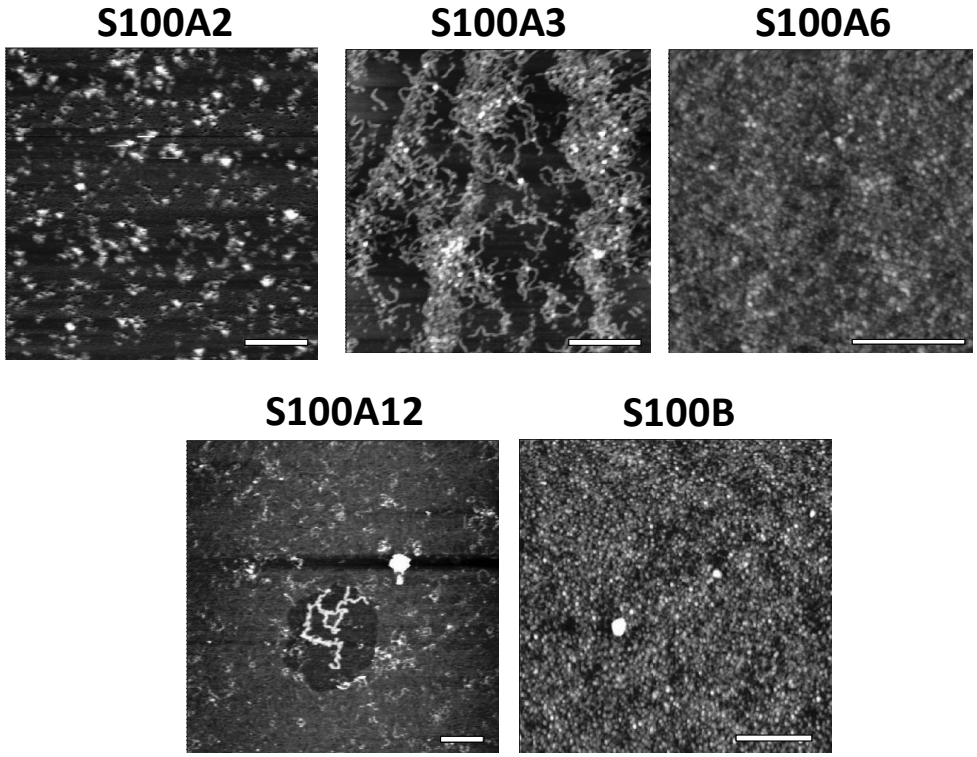


Figure 9.4 – Atomic Force Microscopy analysis of S100 amyloid morphology.

Surprisingly, despite the similar amyloid formation kinetics, all S100 proteins form quite distinct amyloid structures. S100A2 and S100A3, which had a high initial ThT fluorescence exhibit dissimilar structures: S100A2 presents an oligomeric/protofibrillar morphology and S100A3 is highly amyloidogenic, forming proper amyloid fibrils. S100A6 and S100B, which exhibit similar amyloid formation kinetics, rather than forming fibrillar structures, assemble in β-sheet rich oligomeric structures. Most striking is the morphology of S100A12. This protein did not bind ThT even after prolonged incubation at amyloid formation prone conditions but assembled in thin protofilaments. This indicates that despite the microscopic morphological resemblance to amyloid species, the ultrastructural details are distinct, with lower β-sheet content. This morphological analysis indicates that the conformational plasticity of S100 proteins is further evidenced by their amyloidogenic properties, leading to the formation of remarkably different structures despite sequence similarity within the same protein family.

9.4.4. Structural characterization of S100B amyloidogenesis

To gain further structural knowledge of S100 amyloidogenesis we have selected S100B, one of the most abundant proteins in the brain and thus naturally present in high concentration, which is one of the prerequisites for amyloidogenesis. Further, this protein is secreted at micromolar concentrations under particular conditions [23]. To this purpose, we have monitored S100B amyloid formation by Attenuated Total Reflectance Fourier Transform-Infra Red spectroscopy (ATR FT-IR), a particularly adequate technique for describing amyloidogenic processes due to its high intrinsic sensitivity to β structures (Figure 9.5). Due to sensitivity requirements, for this study we have increased the protein concentration to 10 mg/ml (*versus* 3 mg/ml in the ThT-derived amyloidogenesis kinetics), keeping all remaining experimental parameters unchanged.

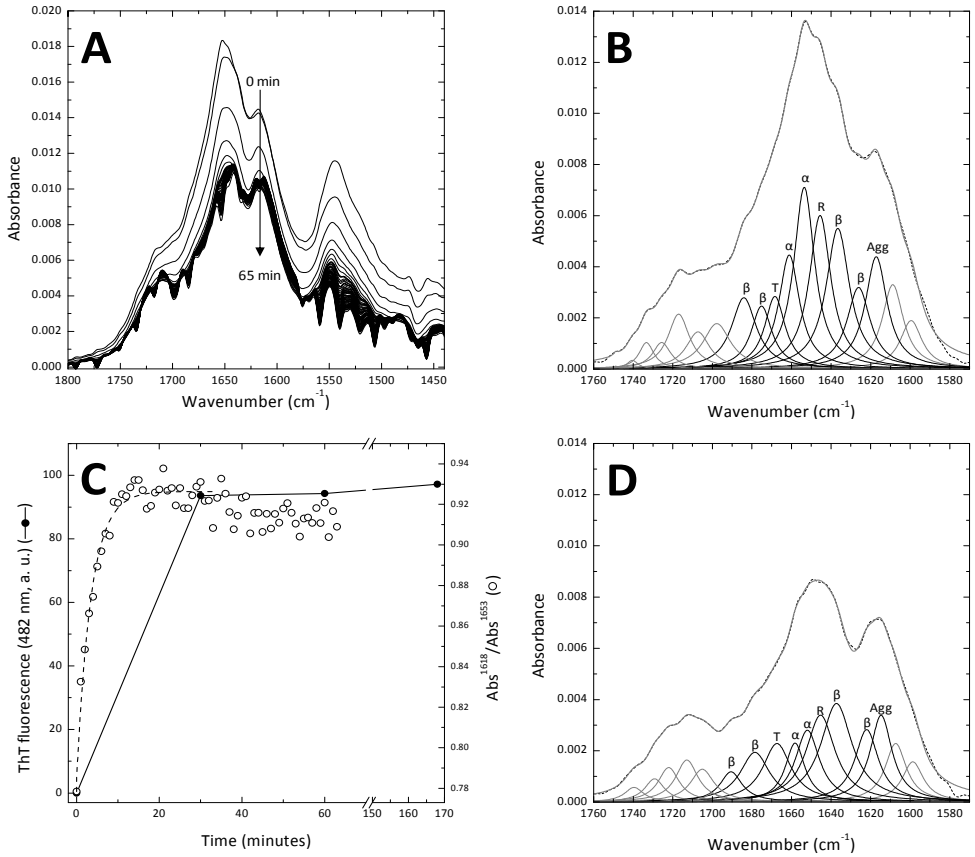


Figure 9.5 – Structural characterization of S100B amyloid formation. (A) FT-IR spectra in the amide I and amide II regions obtained throughout S100B amyloidogenesis. Deconvoluted FT-IR spectra obtained before (B) and after a 6 min incubation (D) at 57°C. Component band assignment is shown in the plots: α-helices (α), β-sheets (β), β-turns (T), random coil (R) and intermolecular aggregation β-sheets (Agg). The FT-IR spectrum is shown as black dashes. The sum of the component bands is shown as a continuous grey line. (C) S100B amyloid formation kinetics monitored through the ThT fluorescence assay and absorption enhancement at FT-IR intermolecular β-sheet characteristic 1618 cm^{-1} wavenumber normalized by the amide I absorption maximum. Experimental points are fitted to a monoexponential decay equation. Protein concentration was 10 mg/ml in 50 mM glycine pH 2.5 and amyloid formation was promoted by incubating at 57°C.

The spectral maximum around 1650 cm^{-1} is characteristic of S100 proteins [24] for their high helical content. The spectra obtained before incubation at the amyloidogenic temperature (57°C) also exhibits a local maximum at 1618 cm^{-1} (Figure 9.5A), which is characteristic of intermolecular β-sheets, occurring in aggregation processes. This indicates that acidic pH alone can induce some S100B aggregation. During incubation

at the amyloidogenic temperature, the relative contribution of the 1618 cm^{-1} band to the overall absorption increases rapidly, revealing amyloid formation, which is well described as a monoexponential decay with a rate constant of 3.51 min^{-1} (Figure 9.5C). Probing the same process by the ThT fluorescence assay yields similar results. We have taken advantage of the quantitative absorption of specific secondary structure types in the amide I spectral region (1600-1700 cm^{-1}) to deconvolute the FT-IR spectra before thermal incubation and at the beginning of the kinetic's plateau phase into individual component bands (Figure 9.5BD). We have assigned each band to the secondary structure element absorbing at each specific wavenumber and quantified structural content by curve integration (Table 9.1).

By comparing the content of each type of secondary structure at the two time points one can assess the conformational changes occurring (Table 9.2). The low helical content is a consequence of solution acidification, which is used to facilitate amyloid generation by enhancing the intrinsic protein aggregation propensity. Throughout amyloid generation 11% of the α-helical and coil content is converted to β-sheets and β-turns, a conformational change typical of amyloidogenic processes. The end state is thus β-sheet rich, as the ThT fluorescence assays also indicate.

Table 9.1 – Assignment of component bands from FT-IR spectra deconvolution.

	0 min		**6 min**	
Structure Type	Wavenumber (cm^{-1})	% Structure	Wavenumber (cm^{-1})	% Structure
α-helix	1654	17	1652	11
	1661	11	1658	7
β-sheet	1626	9	1622	11
	1637	15	1637	19
	1675	6	1678	9
	1684	8	1691	4
B-turns	1668	7	1667	10
Random	1646	15	1645	16
Aggregation sheets	1617	11	1615	12

Table 9.2 – Secondary structure changes occurring during S100B amyloidogenesis. From FT-IR spectra decomvolution.

	% Structure		
	0 min	**6 min**	**Variation (%)**
α-helices	28	18	-10
β-sheets	38	44	+6
β-turns	7	10	+3
Random coil	15	16	-1
Intermolecular β-sheets	11	12	+1

9.4.5. Folding and stability of S100A6

Our amyloid formation screening is based on the observation that the amyloid state is accessible to multiple protein folds [25] and may be attained *via* a destabilized, slightly misfolded native-like state [26]. This observation is crucial to the understanding of pathology-related amyloidogenic properties, since the extensive conformational rearrangement occurring in the proteins involved are thus more likely to happen in the mild physiological milieu. To assess the structural and stability changes occurring in S100 proteins upon lowering the pH we have selected S100A6 as a model. From all the S100 proteins included in our screening, this is the only one whose folding properties were not yet addressed in the literature, so this approach further contributes to the biophysical characterization of S100 protein folding. We have performed thermal denaturation assays on S100A6 placed either on 50 mM KPi pH 7 or 50 mM glycine pH 2.5 while monitoring the far UV CD spectral signal (Figure 9.6).

Like other S100 proteins, the thermal unfolding of S100A6 is irreversible. At pH 7 (Figure 9.6A) S100A6 is rather stable but the high T_m^{app} precludes the detection of the unfolding post-transition. From the accessible temperature range, the T_m^{app} was estimated at 81.4°C. Upon acidification the characteristic α-helical spectra is kept (Figure 9.6B), though with lower intensity, revealing a partially misfolded state of S100A6. We have quantified

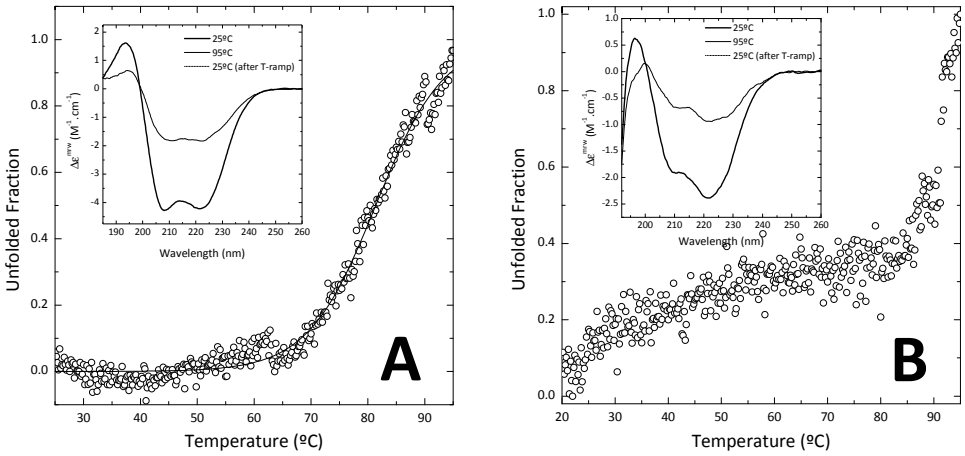


Figure 9.6 – Thermal unfolding of S100A6. Denaturation curves correspond to the normalized CD variation at 222 nm of the apo protein in 50 mM KPi pH 7 (A) or 50 mM glycine pH 2.5 (B). Insets are the far UV CD spectra recorded at 25°C before the temperature ramp, at 95°C and after recooling the sample to 25°C. The sigmoidal fit yields a T_m^{app} = 81.4°C.

this conformational change by deconvoluting the spectra (Table 9.3), estimating the amount of each secondary structural element in this protein. This further indicates that at pH 2.5 there is a minor conversion of helical and turn elements to β-sheets and random coil.

However, acidification has significant effects in the stability of S100A6. Unfolding proceeds through a multistate mechanism, with two clearly distinguishable transitions, the first being uncooperative and occurring around 30°C and the second occurring around 95°C in a highly cooperative manner. The T_m^{app} of this second transition could not be determined

Table 9.3 - Deconvolution of the far UV CD spectra of S100A6 at pH 7 and 2.5 before thermal perturbation.

	Structural content (%)		
	pH 7.0	**pH 2.5**	**Variation (%)**
α-helices	24.1	20.3	-3.8
Antiparallel β-sheets	13.6	17.3	+3.7
Parallel β-sheets	6.1	6.1	-0.1
β-turn	17.5	16.4	-1.2
Random coil	38.6	39.9	+1.4

accurately due to the lack of the post-transition zone. According to these measurements, the incubation at 57°C in the amyloid screening assay yielded a misfolded state, though with residual structure. The fast amyloid formation kinetics could be then attributed to the helical-to-sheet transition brought about by acidification alone, raising the possibility that amyloid formation may be readily accessible at higher pH values.

9.5. Discussion

Even before of the association of S100 proteins with amyloidogenic processes in the prostate it was known that that several S100 proteins accumulate in the brain in benign *corpora amylacea* [27], the etiology of which is unknown and the possible implications in brain inflammatory or other pathologies still unaddressed. On the other hand, the control of S100 amyloid deposition is related to the development of prostate adenocarcinoma wherein prostate *corpora amylacea* abundance is lower than in the healthy population [28]. A correlation between prostate S100A8/A9 *corpora amylacea* and cancer onset through inflammation has been suggested on the basis of co-deposition of bacterial proteins and macrophage activation *in situ* [1]. Following these observations and taking into consideration the sequence homology among S100 proteins, the high expression level of some S100 proteins, their frequent extracellular localization and expression mis-regulation in many pathologies we screened several S100 proteins for amyloid formation in acidic conditions and found widespread amyloidogenic propensity associated with morphological polymorphism. We have deliberately chosen acidic pH conditions for our amyloid formation screening because these are known to facilitate the amyloidogenic conversion of native proteins [13], which would otherwise be seldom reached at more physiological conditions through thermal fluctuations [26]. The “rare” formation of amyloids *in vivo* is partially based in this phenomenon. In this sense, low pH enhances the intrinsic

amyloidogenic potential of proteins and is a useful tool in the characterization of amyloidogenesis.

The high amyloid formation propensity assigned to the hydrophobic core-forming helices (H_I and H_{IV}) suggests that S100 dimerization can work as a mechanism to prevent or minimize amyloidogenesis. All members of the S100 family are homo or heterodimers (with the notable exception of S100G, a smaller family member which functions as a calcium buffer rather than as a calcium signal transducer). Unlike helix H_{III}, which undergoes the EH-hand calcium-induced conformational change, helices H_I and H_{IV} maintain their relative position during calcium binding and keep the dimeric state throughout signaling events. The importance of dimerization as a amyloidogenic protection mechanism has been highlighted in the S100A8/A9 heterodimer [1], where the overall aggregation propensity of the individual subunits decreases upon heterodimer formation. Remodeling of the hydrophobic core in the all helical S100 proteins may thus be an important step in the amyloidogenic conversion of S100 proteins.

The first step in amyloid fibril assembly is the formation of a nucleus – the least stable identifiable species on the aggregation pathway – which is elongated through the polymerization of monomeric or oligomeric soluble proteins. Frequently, the nucleation step is rate-limiting in the amyloidogenesis pathway and gives rise to a lag phase preceding amyloid formation. Still, some proteins do not present such delay and promptly form amyloid when placed in permissive conditions. This is the case of Aβ peptide mutants [29-30]. Contrary to the wild type form, D23N-$A\beta_{1-40}$ fibrillization shows no obvious lag period at 100 μM peptide concentration, with a rate constant of 3.77×10^{-3} min^{-1}. For wt-$A\beta_{1-40}$ the rate constant is 1.07×10^{-4} min^{-1} [29]. Despite the experimental conditions in our survey (57°C pH 2.5) are outside the physiological range, they nevertheless reveal an intrinsic amyloid propensity widespread in the S100 protein family. These

experimental conditions have allowed establishing a proof-of-concept for amyloid formation in this protein family, an approach which has been used previously in the literature for distinct proteins (*e. g.* [13]). Further investigations will address amyloid formation under physiological conditions (pH ~7.0, 37°C) and the effects of molecular crowding, which is known to promote nucleation in amyloid formation [31].

From an evolutionary point of view, the proteins included in our amyloid formation screening represent distinct branches of the S100 family (Figure 9.7) which are thought to have evolved through multiple gene duplication events from a common ancestor protein [3].This suggests that, instead of being a peculiarity of a subgroup of S100 proteins, amyloidogenic behavior may be shared by other family members. Such generality is further based in the S100 family sequence analysis, where amyloid propensity was predicted by distinct algorithms.

Concerning *corpora amylacea*, the protein content is in fact only around 4% (w/w) [32-33]. The major component in these inclusions are hexoses (~88% w/w) in a glycogen-like form [34]. The presence of the abundant

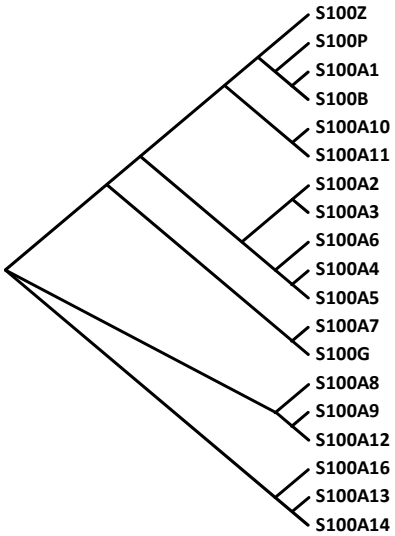


Figure 9.7 – Phylogenetic tree of human S100 proteins. Computed from S100 protein sequences. From [3].

glycan portion may be relevant in the amyloidogenic processes as it is established that macromolecular crowding accelerates the nucleation step of amyloidogenesis and promotes fibrillization of distinct proteins [31]. Further investigations will address the possibility of S100 proteins forming amyloid at physiological-like conditions and the identification of other factors modulating amyloidogenesis. *A priori* candidates are metal ions, macromolecular crowding, oxidant species, amyloid seeds and point mutations, all of which have been associated either with S100 folding or amyloid deposition in general. Also, the correlation of the conditions where amyloid formation is accessible with a more thorough characterization of the structure of amyloid species (oligomers, protofibrils, fibrils) will provide insight on the feasibility of S100 amyloid formation in physiological conditions. Last, the association of S100 amyloids with putative pathologies will require the *in vivo* detection of such species.

By including a comprehensive spectroscopic description of amyloid formation by S100A6 and S100B in our experimental approach, combined with AFM imaging of S100 amyloids as a function of the metallation status, we have indirectly excluded the possibility that the metal-influenced amyloid formation kinetics could be due to the Hoffmeister effect alone: distinct morphologies were observed as a function of metallation.

9.6. Conclusions

Altogether, the above described findings suggest that amyloid-like conformations (β-rich oligomers, protofibrils and fibrils) might be accessible to S100 proteins under particular conditions. The involvement of S100 proteins in several pathologies (including inflammatory processes in the prostate associated to S100A8/A9 amyloids), their deposition in *corpora amylacea* in the healthy brain and the here reported amyloidogenic propensity of several family members points to a new and relevant biomedical research field wherein S100 proteins are hypothesized to form

amyloid deposits associated to several pathologies. The extensively reported overexpression of several S100 proteins in multiple pathologies like cancer, cardiomyopathies, inflammatory conditions and neurodegeneration [3] provides the candidates for putative disorders where S100 fibrillization can be involved. Future research on this field will determine the exact conditions where S100 fibrillization is accessible and if it has pathological relevance apart from the already reported one [1].

9.7. References

1. Yanamandra K, Alexeyev O, Zamotin V, Srivastava V, Shchukarev A, Brorsson AC, Tartaglia GG, Vogl T, Kayed R, Wingsle G, et al. (2009) Amyloid formation by the pro-inflammatory S100A8/A9 proteins in the ageing prostate. *PLoS One* **4**, e5562
2. Fritz G & Heizmann CW (2004) 3D structures of the calcium and zinc binding S100 proteins. In *Handbook of metalloproteins* (Messerschmidt A, Huber R, Poulos T & Wieghardt K, eds). John Wiley & Sons.
3. Marenholz I, Heizmann CW & Fritz G (2004) S100 proteins in mouse and man: from evolution to function and pathology (including an update of the nomenclature). *Biochem Biophys Res Commun* **322**, 1111-1122
4. Guijarro JI, Sunde M, Jones JA, Campbell ID & Dobson CM (1998) Amyloid fibril formation by an SH3 domain. *Proc Natl Acad Sci U S A* **95**, 4224-4228
5. Nerelius C, Fitzen M & Johansson J (2010) Amino acid sequence determinants and molecular chaperones in amyloid fibril formation. *Biochem Biophys Res Commun* **396**, 2-6
6. Fernandez-Escamilla AM, Rousseau F, Schymkowitz J & Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* **22**, 1302-1306
7. Tartaglia GG & Vendruscolo M (2008) The Zyggregator method for predicting protein aggregation propensities. *Chemical Society reviews* **37**, 1395-1401
8. Koch M, Bhattacharya S, Kehl T, Gimona M, Vasak M, Chazin W, Heizmann CW, Kroneck PM & Fritz G (2007) Implications on zinc binding to S100A2. *Biochim Biophys Acta* **1773**, 457-470
9. Tartaglia GG, Pawar AP, Campioni S, Dobson CM, Chiti F & Vendruscolo M (2008) Prediction of aggregation-prone regions in structured proteins. *J Mol Biol* **380**, 425-436
10. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948
11. Nicholas KB, Nicholas HB & Deerfield DW (1997) GeneDoc: Analysis and Visualization of Genetic Variation. *EMBnetnews* **4**, 1-4
12. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254
13. Morozova-Roche LA, Zurdo J, Spencer A, Noppe W, Receveur V, Archer DB, Joniau M & Dobson CM (2000) Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *J Struct Biol* **130**, 339-351
14. Malisauskas M, Zamotin V, Jass J, Noppe W, Dobson CM & Morozova-Roche LA (2003) Amyloid protofilaments from the calcium-binding protein equine lysozyme: formation of ring and linear structures depends on pH and metal ion concentration. *J Mol Biol* **330**, 879-890
15. Barth A & Zscherp C (2002) What vibrations tell us about proteins. *Q Rev Biophys* **35**, 369-430
16. Maurer-Stroh S, Debulpaep M, Kuemmerer N, de la Paz ML, Martins IC, Reumers J, Morris KL, Copland A, Serpell L, Serrano L, et al. (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat Methods* **7**, 237-242
17. Salama I, Malone PS, Mihaimeed F & Jones JL (2008) A review of the S100 proteins in cancer. *Eur J Surg Oncol* **34**, 357-364

18. Goyette J & Geczy CL (2010) Inflammation-associated S100 proteins: new mechanisms that regulate function. *Amino Acids*
19. Komada T, Araki R, Nakatani K, Yada I, Naka M & Tanaka T (1996) Novel specific chemtactic receptor for S100L protein on guinea pig eosinophils. *Biochem Biophys Res Commun* **220**, 871-874
20. Miranda LP, Tao T, Jones A, Chernushevich I, Standing KG, Geczy CL & Alewood PF (2001) Total chemical synthesis and chemotactic activity of human S100A12 (EN-RAGE). *FEBS Lett* **488**, 85-90
21. Strahan B & Arndt P (2010) S100B Regulates Pulmonary Neutrophil Recruitment In Lipopolysaccharide-induced Acute Lung Inflammation. *Am J Respir Crit Care Med* **181**, A2319
22. Leclerc E, Fritz G, Vetter SW & Heizmann CW (2009) Binding of S100 proteins to RAGE: an update. *Biochim Biophys Acta* **1793**, 993-1007
23. Leclerc E, Sturchler E & Vetter SW (2010) The S100B/RAGE Axis in Alzheimer's Disease. *Cardiovasc Psychiatry Neurol* **2010**, 539581
24. Botelho HM, Koch M, Fritz G & Gomes CM (2009) Metal ions modulate the folding and stability of the tumor suppressor protein S100A2. *Febs J* **276**, 1776-1786
25. Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* **75**, 333-366
26. Chiti F & Dobson CM (2009) Amyloid formation by globular proteins under native conditions. *Nat Chem Biol* **5**, 15-22
27. Hoyaux D, Decaestecker C, Heizmann CW, Vogl T, Schafer BW, Salmon I, Kiss R & Pochet R (2000) S100 proteins in Corpora amylacea from normal human brain. *Brain Res* **867**, 280-288
28. Christian JD, Lamm TC, Morrow JF & Bostwick DG (2005) Corpora amylacea in adenocarcinoma of the prostate: incidence and histology within needle core biopsies. *Mod Pathol* **18**, 36-39
29. Tycko R, Sciarretta KL, Orgel JP & Meredith SC (2009) Evidence for novel beta-sheet structures in Iowa mutant beta-amyloid fibrils. *Biochemistry* **48**, 6072-6084
30. Sciarretta KL, Gordon DJ, Petkova AT, Tycko R & Meredith SC (2005) Abeta40-Lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid. *Biochemistry* **44**, 6003-6014
31. Zhou Z, Fan JB, Zhu HL, Shewmaker F, Yan X, Chen X, Chen J, Xiao GF, Guo L & Liang Y (2009) Crowded cell-like environment accelerates the nucleation step of amyloidogenic protein misfolding. *J Biol Chem* **284**, 30148-30158
32. Sakai M, Austin J, Witmer F & Trueb L (1969) Studies of corpora amylacea. I. Isolation and preliminary characterization by chemical and histochemical techniques. *Arch Neurol* **21**, 526-544
33. Steyaert A, Cisse S, Merhi Y, Kalbakji A, Reid N, Gauvreau D & Lacoste-Royal G (1990) Purification and polypeptide composition of corpora amylacea from aged human brain. *J Neurosci Methods* **31**, 59-64
34. Stam FC & Roukema PA (1973) Histochemical and biochemical aspects of corpora amylacea. *Acta Neuropathol* **25**, 95-102

10. Metals as modulators of amyloidogenesis: the example of the S100 family

Part of this chapter was published in

Fritz G, Botelho HM, Morozova-Roche LA, Gomes CM (2010) Natural and amyloid self-assembly of S100 proteins: structural basis of functional diversity. *FEBS J*, 277(22) 4578-4590

Proteins were purified by Gunter Fritz (Freiburg University, DE). AFM analysis was performed by Kiran Yanamandra and Ludmilla A. Morozova-Roche (Umeå University, SE). Cytotoxicity measurements were performed jointly by Hugo M. Botelho, Kiran Yanamandra and Ludmilla A. Morozova-Roche at Umeå University.

10.1. Summary

Amyloid deposition constitutes the cytological hallmark of several diseases, where metal ions are important players, either because they favor fibrillization or act as toxic agents. We have assessed the role of metal ions as modulators of the amyloidogenesis pathways of representative members of the S100 protein family (S1002, S100A3, S100A4, S100A6, S100A12 and S100B). We have observed that S100 amyloid structure is highly sensitive to metallation. Further, in most cases S100 amyloid species are oligomeric. In S100A6 Ca^{2+} binding induces a non-amyloid and non-toxic off-pathway conformation which nevertheless has an increased β-sheet content. The oligomeric nature of most S100 amyloids is associated with cytotoxicity of S100A6 oligomers, S100B mature amyloids and S100A12 oligomers and mature fibrils. These proteins are overexpressed in ALS and Alzheimer's disease and cytotoxicity may be relevant in the understanding of the associated neurodegenerations if S100 fibrillization has physiological relevance.

10.2. Introduction

S100 proteins constitute a family of metal signaling buffers or transducers. At the protein level, metal ions regulate protein conformation and/or oligomerization, therefore determining function. Ca^{2+} binding triggers the characteristic EF-hand conformational change whereby helix H_{III} moves around 90° [1], exposing a hydrophobic patch responsible where downstream signaling partners bind to [2]. Zn^{2+} and Cu^{2+} are usually associated to subtle conformational changes. On the other hand, metal-bound S100 proteins frequently exhibit differential stability. Both Ca^{2+} and Zn^{2+} are S100 oligomerization modulators. Ca^{2+} functions as a dimer cross-linker in the assembly of the S100A8/A9 heterodimer [3], S100A12 hexamer [4] and S100B octamers [5]. Zn^{2+} drives the assembly of the S100A2 tetramer [6].

Moreover, S100A8/A9 oligomerization into amyloid fibrils is dependent of Ca^{2+} or Zn^{2+} *in vitro* [7]. At the cellular level, metal homeostasis and S100 metallation are responsible for determining protein compartmentalization. This exposes S100 proteins to distinct interacting partners and medium compositions. S100B is secreted from astrocytes upon elevation of cytosolic Ca^{2+} concentration or decrease in intracellular Zn^{2+} concentration [8]. In smooth muscle cells S100A1, S100A4 and S100A6 relocate from the sarcoplasmic reticulum to perinuclear vesicles [9].

Our analysis of S100 amyloid-forming propensity revealed that helices H_I and H_{IV}, forming the S100 hydrophobic core, are highly amyloidogenic (Figure 9.1). From our demonstration that S100 proteins form amyloid species (Chapter 9) and the knowledge of the α-helical S100 fold, one may hypothesize that hydrophobic core remodeling yielding enhanced β-sheet content is an early step in S100 amyloidogenesis. In this sense, metal ions can be hypothesized to be important modulators of S100 amyloidogenesis. The calcium-induced conformational change which enhances the solvent exposure of the hydrophobic core could then facilitate S100 amyloidogenesis. Similarly, other metal ions could fulfill similar roles.

Here we have screened for the effect of calcium and zinc in the amyloidogenesis kinetics and amyloid oligomeric structure of S100A2, S100A3, S100A4, S100A6, S100A12 and S100B. We have found a strong influence of metallation on the kinetics and endpoint morphology. Remarkably, S100A6 amyloidogenesis was completely inhibited by Ca^{2+}. A detailed analysis of amyloid formation kinetics by S1006 revealed that Ca^{2+}-S100A6 formed oligomeric structures with increased β-sheet content which nevertheless lacked the amyloid-specific cross-beta core, as shown by the absence of ThT fluorescence. The potential cytotoxicity of S100 oligomers was assessed the quantification of SHSY-5Y neuroblastoma cell viability. S100A6 oligomers, S100A12 oligomers and mature amyloids and S100B

mature amyloids were found to be cytotoxic. Apart from the structural importance, Ca^{2+} is also relevant in modulating the cytotoxic potential of S100 amyloids as it reverts the cytotoxicity of S100A6 and S100B structures, as well as of S100A12 mature amyloids. This is particularly relevant considering that S100 proteins are overexpressed in Amyotrophic Lateral Sclerosis, Alzheimer's and Parkinson's diseases [10-13].

10.3. Materials and methods

10.3.1. Chemicals and proteins

All reagents were of the highest grade commercially available. Thioflavin T was obtained from Sigma. Trace metals were removed from all solutions by solution treatment with a chelex resin. Metal ion solutions were prepared in chelex-treated water. S100 proteins (S100A2, S100A3, S100A4, S100A6, S100A12 and S100B) were expressed in *E. coli* and purified to homogeneity as described previously [6]. Molar protein concentrations refer to the S100 subunit. Hen egg white lysozyme was obtained from Fluka.

10.3.2. S100 fibrillization

S100 proteins were placed in 50 mM Glycine buffer pH 2.5 by repeated dilution and re-concentration. Sample temperature was kept at 4°C at all times before amyloid formation assays. Protein concentration was set at 3 mg/ml using Bradford's method [14]. Due to spontaneous and extensive aggregation in aerobic conditions, S100A3 (4 mg/ml) was previously reduced with 25 mM TCEP overnight. S100 proteins were prepared in the apo state or in the presence of 10 fold molar excess over the S100 monomer (~2.6-2.8 mM) of known metal ion ligands, as the $ZnCl_2$, $CaCl_2$ or $CuSO_4$ forms. Binding was promoted by incubating from 15 min to 2 h at 4°C. Any pre-existing aggregates were pelleted and removed by centrifuging at 12000 g for 10 minutes. Amyloid formation was promoted by quiescent incubation at 57°C. For S100A6 samples containing EDTA, an excess amount was added

at before incubation (apo-S100A6) or at a selected time point (Ca^{2+}-S100A6). The positive control for amyloid formation was hen egg white lysozyme, at 10 mg/ml in the same buffer and incubated at 57°C, a previously validated assay [15].

10.3.3. Thioflavin T fluorescence

ThT binding assays were performed according to [15]. Briefly, 5 µl of protein sample (15 µg) was mixed with 65 µM ThT in 10 mM potassium phosphate (KPi), 150 mM NaCl, pH 7.0 at room temperature to a final volume of 1 ml. Samples were incubated under stirring for 1-2 min and ThT fluorescence at 482 nm was recorded with a Cary Varian Eclipse instrument with temperature set at 25°C using a peltier-thermostated cell support. Excitation wavelength was 440 nm, PMT voltage 600 V, and excitation and emission slits 10 nm.

10.3.4. ATR FT-IR spectroscopy

Secondary structural changes occurring during S100A6 amyloid formation were monitored by ATR FT-IR measurements performed in a Bruker IFS 66/S spectrometer equipped with a nitrogen-cooled MCT detector using the thermostatized Harrick BioATR II cell. 3 mg/ml S100A6 in 50 mM glycine pH 2.5 were prepared in the apo state, in the apo state with excess EDTA and in the presence of mM Ca^{2+} (*i. e.* 10 molar equivalents). Samples were incubated at 57°C while spectra were continuously acquired (3 min accumulation time, 12 mm aperture, 20 kHz scanner velocity). Spectral resolution was 4 cm^{-1}. Amyloid formation was assessed by the absorption ratio between the intramolecular β-sheet characteristic band at 1626 cm^{-1} and the amide I absorption maximum, 1655 cm^{-1}.

10.3.5. Atomic force microscopy

After reaching the plateau phase in amyloid formation kinetics, S100 samples were stored at 4°C and morphology was analyzed by AFM on a PICO PLUS 5500 microscope (Agilent) as previously described [16].

10.3.6. Dynamic light scattering

For monitoring molecular size variations during amyloid formation of S100A6, 3 mg/ml protein solutions were prepared in the apo state or in the presence of 10 molar equivalents (2.9 mM) of Ca^{2+} or Zn^{2+} in the form of chloride metal salts in 50 mM glycine pH 2.5. Samples were then filtered through a 0.22 µM membrane and analyzed with a Malvern Instruments Zetasizer Nano ZS instrument equipped with a 633 nm laser while the temperature was kept at 57°C by a Peltier-controlled thermostatized cell support. Kinetic data was obtained from continuously repeated measurements. In each measurement, the backscattered light (173°) from fifteen 10 s accumulations was averaged. Amyloid formation was assessed from light scattering intensity and the average particle size (diameter) obtained from the intensity distribution. Results were analyzed with Malvern Instruments DTS software using a multimodal fit with quadratic weighting and 0.01 regulariser.

10.3.7. Amyloid cytotoxicity

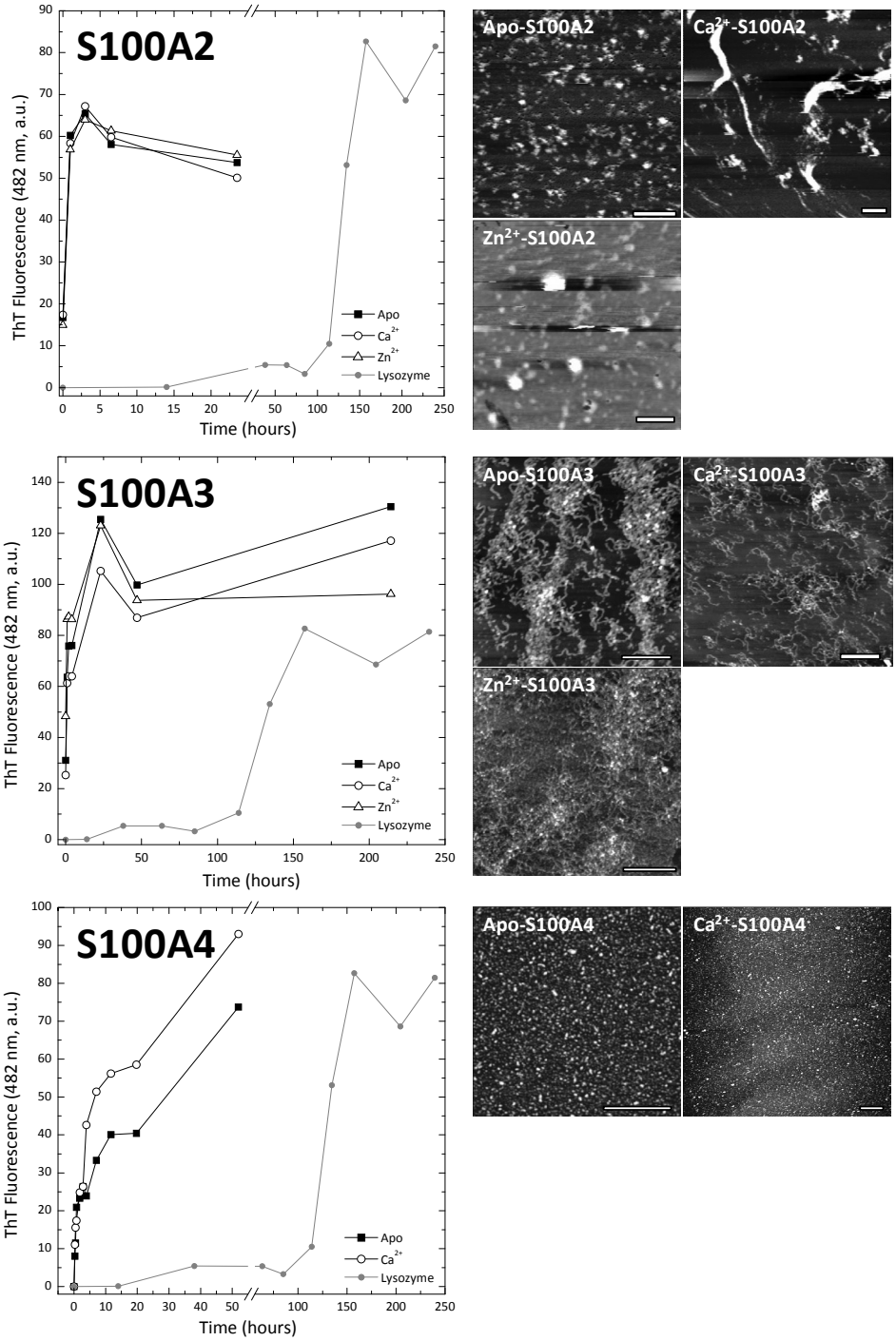
For cytotoxicity measurements S100A6, S100A12 and S100B samples were prepared at 3 mg/ml in 50 mM glycine pH 2.5 and incubated at 57°C. Distinct samples were prepared in the presence of 10 molar equivalents relatively to the S100 subunit. Several species in the amyloid formation pathway could be obtained by varying the incubation time: native protein (no incubation), oligomeric structures (0.5 or 1 h incubation for S100A6; 2 h for S100A12 and S100B) or endpoint structures (~3 days incubation). S100 cytotoxicity towards the human neuroblastoma SH-SY5Y cell line was

assessed using the WST-1 assay. In viable cells, WST-1 is reduced by mitochondrial reductases, forming formazan dyes, which serves as an indicator of the amount of metabolically active cells. The amount of WST-1 can be quantified by its characteristic absorption around 450 nm [17]. Cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% (v/v) fetal bovine serum and antibiotics in a 5% CO_2 humidified atmosphere at 37 °C. Measurements were performed on cells seeded overnight in 96 well plates (10^4 cells/well). Before starting the measurement, culture medium was replaced by 100 µl fresh medium without fetal bovine serum and S100 samples were added. Cytotoxicity was determined 24 and 48 h after placing S100 samples by adding 10 µl WST-1 reagent solution (Roche) at each well. Cells were further incubated for 4 h and absorbance was measured with an ELISA plate reader (Labsystem Multiscan RC) at 450 nm after cell shaking. Cell viability was expressed as the absorption percentage relatively to control cells which were incubated with buffer alone (50 mM glycine pH 2.5). Statistical differences were assessed with double-tailed, two sample unequal variance Student's *t*-test. For metallated S100 samples, additional control wells were prepared which did not contain protein but metal ions alone at the same concentration.

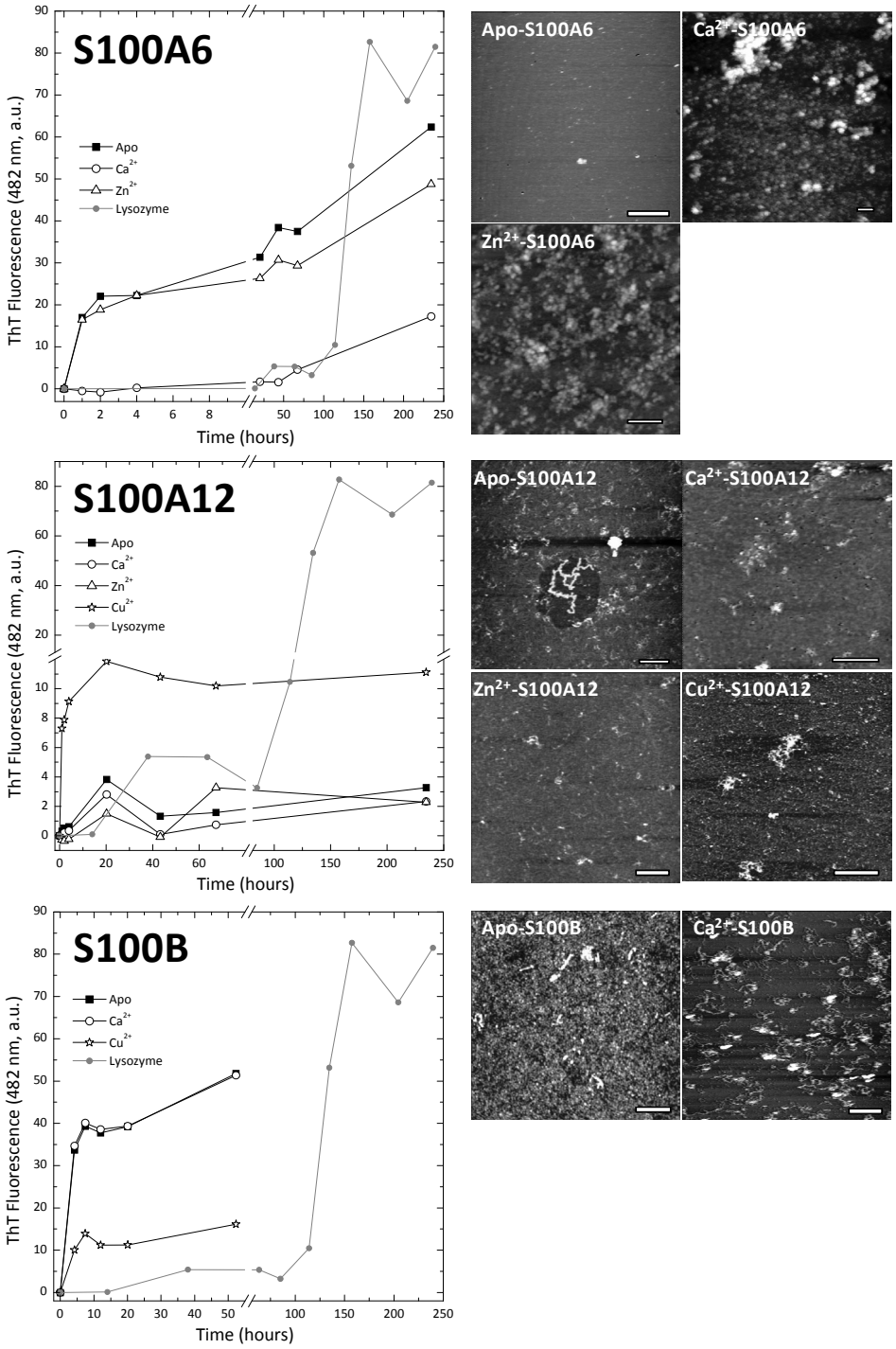
10.4. Results

10.4.1. Metal ions modulate S100 amyloidogenesis

To determine the effect of metallation in S100 amyloid formation kinetics, we have built on our observation of amyloid formation by several S100 proteins at acidic pH (Chapter 9). We have prepared S100 samples (S100A2, S100A3, S100A4, S100A6, S100A12 and S100B) as in the previous assay (3 mg/ml, 50 mM glycine pH 2.5) and included 10 molar equivalents of known metal ion ligands. Samples were incubated quiescently at 57°C and time



(Legend on page 251)



(Legend on page 251)

Figure 10.1 – Metal ions and S100 amyloidogenesis. (Left) Amyloid formation kinetics of S100A2, S100A3, S100A4, S100A6, S100A12 and S100B at 57°C as assessed by the thioflavin T fluorescence assay. S100 proteins were prepared in 50 mM glycine pH 2.5 in the apo state or in the presence of 10 fold molar excess of Ca^{2+}, Zn^{2+} or Cu^{2+}. Lysozyme is the positive control for amyloid formation. (Left) AFM images of protein samples after the kinetic assay. Scale bar: 1 μm.

resolved amyloid formation was quantified using the Thioflavin T (ThT) fluorescence assay. In amyloidogenesis kinetics after a plateau was reached, the morphology of endpoint structures was analyzed by atomic force microscopy (AFM) (Figure 10.1). This complementary approach provides information about the overall secondary (ThT) and quaternary structure (AFM) of S100 oligomers. However, AFM is not a quantitative technique, precluding such kind of image analysis. Data from apoprotein preparations was used as reference.

S100A2 amyloid formation kinetics are unaffected by metal load. All S100A2 preparations show enhanced fluorescence before incubation, an indication that low pH alone can induce an amyloid-like state and reach near final signal values after ~1 h, an indication that low pH alone is able to induce a β-sheet rich conformation in S100A2, suggesting an amyloidogenic potential of this protein, in line with theoretical calculations (Figure 9.2). Whereas apo-S100A2 amyloid species are protofibrilllar, the presence of Ca^{2+} or Zn^{2+} results in the formation of large oligomeric assemblies, which are irregularly shaped in the presence of Ca^{2+} or spherical in the presence of Zn^{2+}. Specifically for S100A2, the starting point for the Zn^{2+} sample should be the homotetramer [6] because of the metal ion concentration (2.7 mM, 10 fold excess. However, this must not be the only factor determining the distinct morphology of Ca^{2+}- and Zn^{2+}-S100A2 as the binding sites for each metal are distinct [6].

Like S100A2, S100A3 seems to rest in a β-sheet enriched state at low pH, as seen by the high ThT fluorescence before the amyloid formation assay. ThT-derived amyloidogenesis kinetics is similar for the three metallated

states. Fluorescence intensity at the plateau phase is higher than the standard lysozyme curve, which was obtained with a higher protein concentration (10 mg/ml lysozyme *versus* 3 mg/ml S100A2). The reason for this is that S100A3 is highly amyloidogenic, forming long amyloid fibrils in all metallated states.

S100A4 forms ThT-reactive species independently of the presence of Ca^{2+} and with a similar kinetics. Nevertheless, structures of apo- and Ca^{2+}-S100A4 amyloid species are distinct, as assessed by the higher ThT fluorescence emission in the latter. The morphological analysis of Ca^{2+}-S100A4 endpoint structures revealed small oligomers.

S100A6 exhibited an extreme case of metal-modulated amyloidogenesis. Apo-S100A6 promptly formed ThT-reactive species, the process not being affected by the presence of Zn^{2+}. Notably, Ca^{2+} completely inhibited the process. However, after extensive incubation, all S100A6 species are of the oligomeric type, irrespectively of metal load. Both Ca^{2+} and Zn^{2+} induced S100A6 aggregation into large structures, unlike the small oligomers seen in the apo protein preparation. The amyloid formation kinetics and the ultrastructural details of the different metallated forms of S100A6 were addressed more thoroughly in Section 10.4.2.

In S100A12, metal ions also exerted a significant influence in the amyloidogenic process. Neither apo-, Ca^{2+}- nor Zn^{2+}-S100A12 formed amyloid species. However, Cu^{2+}-S100A12 promptly assembled in amyloid species, which nevertheless exhibited lower ThT fluorescence than other S100 ThT-reactive preparations. The AFM analysis revealed that all metallated S100A12 forms formed oligomeric species. In the Zn^{2+}-S100A12 sample some protofibrils were also visible.

Regarding S100B, amyloid species were formed from the apo and all metallated forms. Cu^{2+}-S100B exhibited a lower fluorescence intensity, which suggests lower amyloid formation efficiency or an alternative conformation.

In the presence of Ca^{2+} S100B forms oligomeric and protofibrilllar assemblies, unlike the small oligomers originating in the apo protein preparation.

10.4.2. Reversible control of S100A6 amyloidogenesis by Ca^{2+}

The distinct roles of metal ions in S100A6 amyloidogenesis – Zn^{2+} affecting morphology without modifying the formation kinetics and Ca^{2+} inhibiting amyloid formation – have prompted us to investigate this aspect in more detail. To determine the reversibility of the inhibition of S100A6 amyloidogenesis by Ca^{2+} we placed a Ca^{2+}-S100A6 sample in amyloidogenic conditions for 3 h, at which time an excess of EDTA was added (Figure 10.2A). Promptly, ThT-positive species built up, with a plateau ThT fluorescence similar to the one of S100A6 incubated with EDTA alone. Moreover, at the end of incubation the AFM-assessed morphology of EDTA-S100A6 and Ca^{2+}-EDTA-S100A6 was of the oligomeric type, similar to the one of apo-S100A6 but unlike the ones obtained in the presence of Ca^{2+} alone, which were much larger. These results indicate that amyloid formation by S100A6 can be controlled – even at low pH and moderate temperature – by the availability of Ca^{2+}.

Because amyloid formation is a process wherein β-sheet content increases, Fourier transform infrared spectroscopy (FT-IR) is one of the best suited technique to study it, due to its high sensitivity to such structural features [18]. FT-IR absorption at the amide I band (1600-1700 cm^{-1}) arises from distinct secondary structure elements in a quantitative manner, implying that absorbance changes can be assigned to specific conformational changes. Of particular interest to fibrillization processes is the fact that characteristic absorption peaks for α-helices (~1654 cm^{-1}), intramolecular β-sheets (~1684 and 1633 cm^{-1}) and intermolecular β-sheets (~1620 cm^{-1}) are well established [18]. In particular, intermolecular β-sheets are characteristic of amyloid species. To dissect the mechanism of Ca^{2+}-induced inhibition of

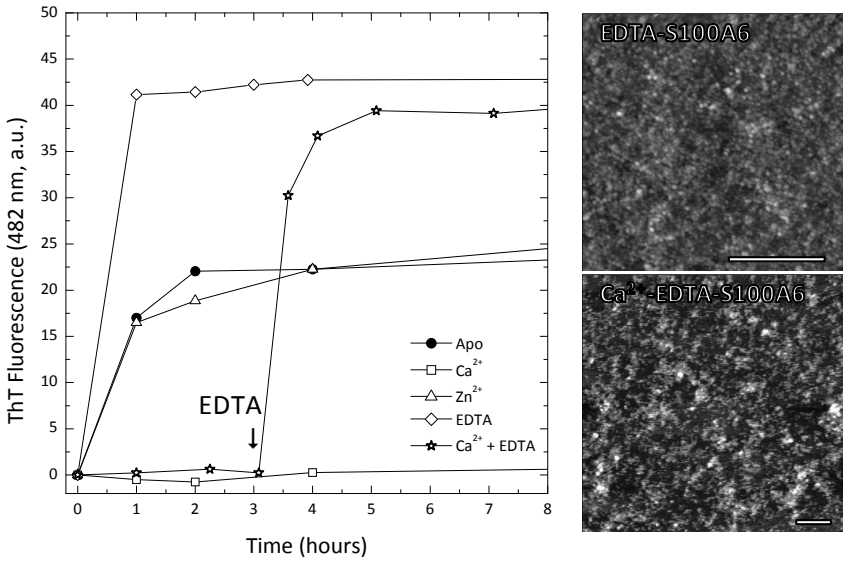


Figure 10.1 – Modulation of S100A6 amyloidogenesis kinetics by Ca^{2+}. (Left) S100A6 amyloid formation kinetics at 57°C assessed by the thioflavin T fluorescence assay as [15]. For analyzing the amyloidogenesis suppression by Ca^{2+}, excess EDTA was added to apo-S100A6 before incubation (✯) or to Ca^{2+}-S100A6 after 3 h incubation (◇). (Right) AFM pictures of the EDTA-S100A6 and Ca^{2+}-EDTA-S100A6 obtained after the kinetic's plateau phase. Scale bar: 1 µm.

S100A6 fibrillization, we have prepared apo-S100A6, EDTA-S100A6 and Ca^{2+}-S100A6 samples as before and followed amyloidogenesis by FT-IR (Figure 10.3).

We observed that amyloid formation was accompanied by the development of a new spectral band at 1626 cm^{-1} which, by its intermediate position, we have assigned to a mixture of intra and intermolecular β-sheets. The amide I absorption maximum occurs at 1655 cm^{-1}, a typical value for S100 proteins [19] due to the α-helical fold. We have expressed the amyloidogenesis kinetics by the absorbance ratio between these two wavenumbers (Figure 10.3D). FT-IR data are adequately described as single or double exponential decays. In accordance, we computed kinetic rate constants which describe the amyloidogenic processes (Table 10.1). Apo- and EDTA-S100A6 traces were virtually superimposable in the first 1.5 h. Apo-S100A6 amyloidogenesis kinetics was double exponential, with rate

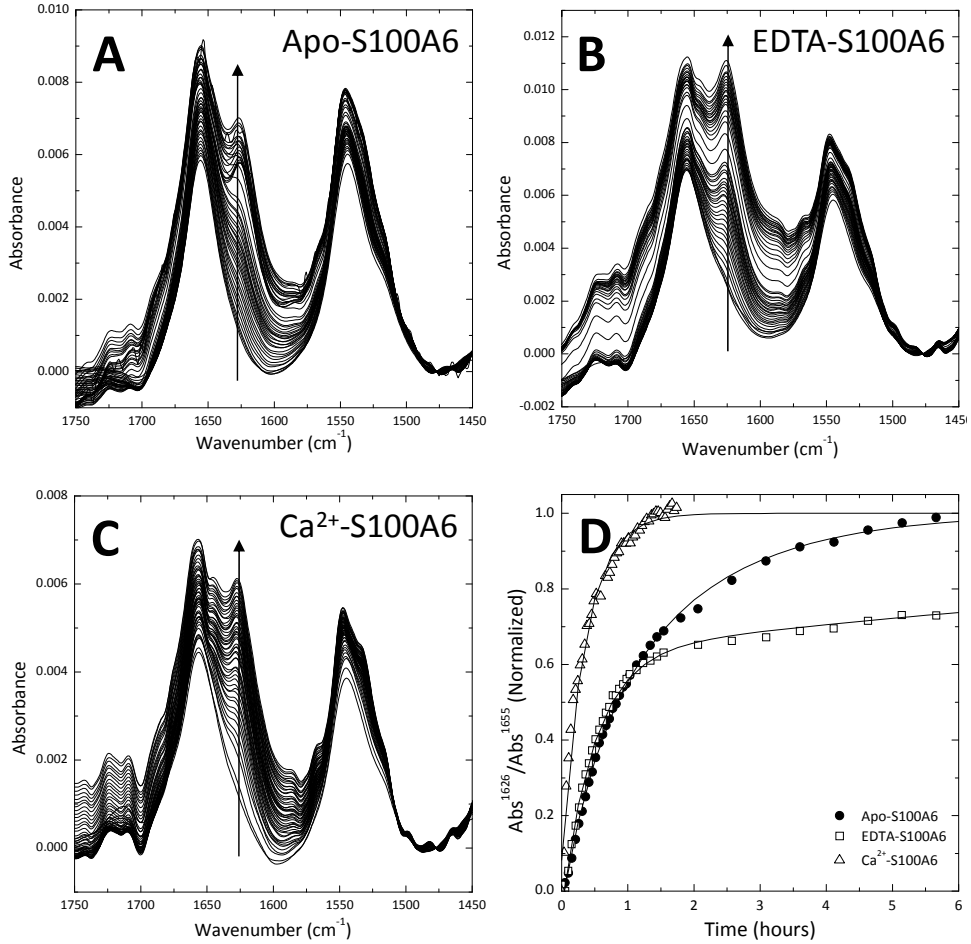


Figure 10.3 – ATR FT-IR monitored amyloid formation by S100A6. Spectral changes in the amide I/II bands during incubation of 3 mg/ml S100A6 in 50 mM glycine pH 2.5 in the apo state (A), in the apo state with excess EDTA (B) and in the presence of 2.9 mM Ca^{2+} (C). (D) Kinetic traces correspond to the absorbance variation at 1626 cm^{-1} (associated with intramolecular β-sheets) normalized by the absorbance at 1655 cm^{-1} (spectral maximum).

constants 0.51 and 1.71 h^{-1}. The fastest rate constant is similar to the one obtained for the single decay of EDTA-S100A6 (0.36 h^{-1}), suggesting that the slowest process could be due to a small amount of contaminant metal ions. For Ca^{2+}-S100A6, the FT-IR derived kinetics was overall faster.

The other characteristic biophysical feature of amyloid formation apart from β-sheet accumulation is protein particle size increase. We have assessed size variation of apo- and Ca^{2+}-S100A6 in real time by dynamic light

scattering (Figure 10.4). Apart from the increasing light scattering intensity due to aggregation processes, we have also analyzed the particle size obtained through autocorrelation analysis of Brownian light scattering intensity fluctuations, which complements AFM measurements with time-resolved size information.

The kinetic information extracted from all techniques is similar, reporting full amyloid conversion around 10 h of incubation in amyloidogenic conditions. Light scattering intensity yields slower kinetics. This occurs because light scattering is proportional to the inverse square of particle size. Consequently, light scattering intensity is biased to higher particle sizes like the ones of endpoint S100A6 structures. We observed that amyloidogenesis rate constants obtained through different techniques (Table 10.1) are similar for apo- and Ca^{2+}-S100A6. This means that Ca^{2+} does not inhibit amyloid formation by S100A6 by locking it in a native-like conformation but by inducing an alternate aggregated, β-sheet rich conformation off the amyloidogenic pathway. The different morphology can be appreciated by the different oligomer size for apo-S100A6 (~14.5 nm) and Ca^{2+}-S100A6 (15.5 nm).

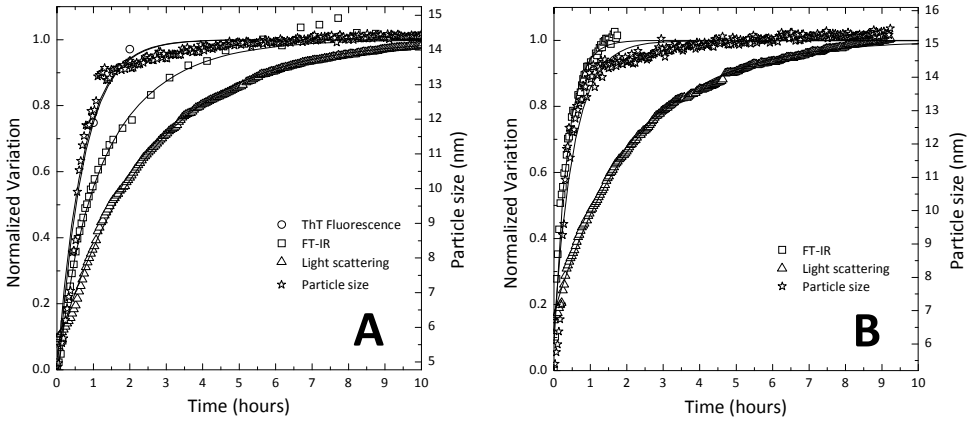


Figure 10.4 - S100A6 amyloidogenesis kinetics. Kinetic traces obtained through ThT fluorescence at 482 nm, FT-IR absorbance change at 1626 cm^{-1}, light scattering intensity and particle size variation. Data points were fitted to mono or double exponential decay equations. (A) Apo-S100A6. (B) Ca^{2+}-S100A6.

Table 10.1 - Rate constants for S100A6 amyloidogenesis. n/a: not applicable, Ca^{2+}-S100A6 does not form ThT-reactive species. n.d.: not determined. *Values for ThT fluorescence-derived kinetics are estimations due to the low data point number.

	Rate constants (h^{-1})		
	Apo-S100A6	**Ca^{2+}-S100A6**	**EDTA-S100A6**
ThT fluorescence*	0.70	n/a	n.d.
FT-IR	0.51 1.71	0.50 17.29	0.36
Light scattering intensity	2.55	2.24	n.d.
Particle size	0.69	0.56	n.d.

10.4.3. Cytotoxicity of S100 amyloid structures

Noteworthy, most S100 amyloid structures are of the oligomeric type. Oligomers have been proposed as a general cytotoxic species [20]. Specifically, in Alzheimer's disease, cognitive impairment is associated with oligomeric Aβ buildup than with senile plaque formation. Accordingly, it is relevant to examine the toxicity of S100 amyloid oligomers. Due to the specific cell expression pattern of S100 proteins, we have selected a subgroup composed of S100A6, S100A12 and S100B to examine their cytotoxic effects. All these proteins are expressed in the brain, some of them at high levels (S100B represents 0.5% of total brain protein and S100A12 constitutes about 5% of total cytosolic protein in resting neutrophils [21]) and are involved in amyloid neurodegenerative diseases: S100A6 is overexpressed in Amyotrophic Lateral Sclerosis (ALS) [11] and Alzheimer's disease [10]; S100A12 and S100B are overexpressed in Alzheimer's disease [12-13]. To assess cytotoxic effects we have selected the human neuroblastoma SHSY-5Y cell line, a well established model for these studies. We have quantified cell viability in the presence of S100 native and amyloid species – oligomers obtained at early stages of the amyloid formation pathway and endpoint structures – by measuring the reduction of the artificial chromogenic electron acceptor WST-1. To examine the role of metal ions in modulating S100 cytotoxicity, we have analyzed S100 species in the

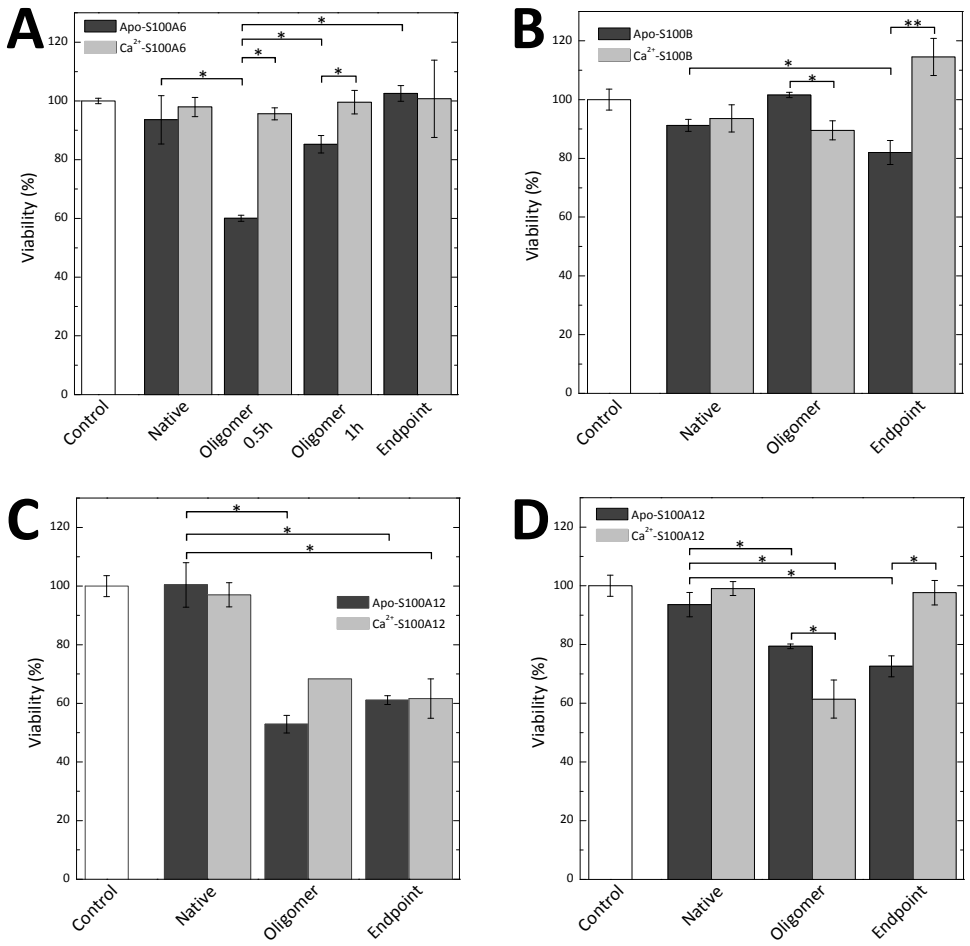


Figure 10.5 – Cytotoxicity of native and amyloid S100 species. S100 proteins at distinct points in the amyloid formation pathway (57°C, pH 2.5) in the apo or Ca^{2+}-bound states were added to SHSY-5Y cell cultures and cellular viability relatively to unperturbed controls was assessed by the WST-1 reduction assay. (A) Cytotoxicity of 5 μM S100A6 species after 48 h incubation. (B) Cytotoxicity of 5 μM S100B species after 24 h incubation. (C-D) Cytotoxicity of S100A12 species after 24 h incubation at 5 μM (C) or 25 μM (D) protein concentration. Error bars represent the standard deviation (n=3).*: $p<0.05$. **: $p<0.1$.

apo and calcium-bound states (Figure 10.5). Zinc and copper were excluded from our screening because of their inherent toxicity at the concentration range for this assay. Protein concentration and incubation time were selected so that the native protein was not toxic itself. In the sole presence of buffer or calcium at the maximal tested concentration, cell viability was identical to unchallenged controls after 24 and 48 h (not shown).

Apo-S100A6 oligomers were found to be cytotoxic but toxicity decreased as higher order oligomeric structures formed at the expense of these (Figure 10.5A). Apart from inhibiting amyloid formation, Ca^{2+} also induced a non-toxic conformation, effectively reverting toxicity. S100B mature amyloid species exhibited low cytotoxicity (~80% cell viability *versus* control, Figure 10.5B), but the remaining species were not significantly toxic irrespectively of metal load. The cytotoxicity profile of S100A12 (Figure 10.5C) was concentration dependent. At 5 μM, the oligomeric and endpoint structures were equally toxic (50-60% cell viability), independently of metal load. At 25 μM, oligomeric and mature apo-S100A12 species were cytotoxic. Ca^{2+}-S100A12 oligomers exhibited enhanced toxicity, which is fully abolished in as mature species form. Interestingly, no ThT binding was observed for apo- of Ca^{2+}-S100A12. Differential toxicity then reveals that incubation at amyloid prone conditions is responsible for a conformational transition distinct from amyloid β-sheet buildup. The protofibrilllar structures of the mature apo-S100A12 misfolded conformation compared to the oligomers formed by Ca^{2+}-S100A12 (Figure 10.1) is probably the molecular basis for the differential toxicity.

10.5. Discussion

The results presented herein constitute novel evidence bridging S100 conformational control, misfolding and molecular neurobiology. For this study we have chosen S100 proteins representative of different family evolutionary branches to reveal the potential for metal ion modulation of S100 amyloidogenic processes. Our results show that despite the amyloidogenic potential being shared by all proteins, the actual structure is frequently oligomeric rather than fibrillar. S100A3 is an exception in the sense that it forms archetypal amyloid fibrils. In almost all proteins analyzed Ca^{2+} promotes the formation of ticker aggregate structures.

Amyloid oligomers, *i. e.* oligomeric structures with enhanced β-sheet content, are intermediates in the amyloid formation pathways where they precede fibril formation and disappear after fibrils have formed. Despite sharing the β-sheet component which is responsible for ThT intercalation and fluorescence, they are most unlike short fibrils. In fact, amyloid oligomers from different proteins share morphological features which can be identified using conformation-sensitive antibodies [22]. They are relevant in neurodegenerative processes like Alzheimer's disease, where disease severity correlates better with the amount of oligomeric Aβ than with the amount of fibrillar deposits [23-24].

Our cytotoxicity assessment was focused in brain proteins overexpressed in neurodegenerative diseases because toxicity has already been described in such cases the origin of which is not yet fully established. Moreover, S100 proteins are present in *corpora amylacea* in the normal brain [25] and these structures have been associated with S100 amyloidogenic conversion in the prostate, involving locally abundant calcium, iron and copper [7]. In Alzheimer's disease (AD), Aβ hexamers or tetramers are thought to insert in membranes [26] forming pores [27-29] which disturb calcium homeostasis [30]. Larger species are less toxic because inter-subunit contacts outcompete peptide-membrane interactions [26]. It is not clear at the moment if a similar mechanism may apply to S100 amyloids. Such determination will require a thorough analysis of the structure of S100 oligomers' structure as well as the consideration of other cytotoxicity mechanisms (*e. g.* via RAGE activation). Nevertheless, in the presence of calcium, the S100A12 homodimer may spontaneously form hexamers and S100B is known to assemble into octameric structures. As our results further support, amyloid-associated toxicity is strongly associated to oligomeric-type structures [20]. Indeed, even after prolonged incubation in amyloidogenic conditions, most S100 protein preparations exhibit an oligomer-like morphology. It is possible that

toxicity is detected in some endpoint preparations because of residual amounts of more toxic oligomeric-like species. However, the amounts of such oligomers should be low as all endpoint preparations were collected at the plateau phase of the ThT fluorescence-monitored amyloid formation kinetics trace.

Toxicity induced by S100 proteins is established in the literature. S100B is known for its concentration-dependent cellular effects in the central nervous system, being neurotrophic at nanomolar levels [31] and inducing apoptosis in a RAGE-dependent manner [32] and exacerbating Aβ neurotoxicity [33] at micromolar concentrations. The novelty in our results is the description of increased toxicity associated solely with the amyloid conversion of S100 proteins.

Metal ion binding affects the protein folding energy landscape. Regarding amyloid formation, metal ions may facilitate the conformational change resulting in the amyloid state, stabilize partially unfolded and amyloid-prone states or stabilize the amyloid conformation. This cross-linking effect is akin to a covalent modification. An artificial cross-linking between Asp23 and Lys28 side chains, stabilizing the β-turn in the amyloid conformation of the $A\beta_{1\text{-}40}$ peptide eliminated the lag phase present in the wild type peptide and increased the formation rate constant 1000 fold [34], suggesting that local conformational constraining accelerates fibril nucleation or fibril growth. On the other hand, zinc promptly induces Aβ fibrillization and stabilizes toxic oligomeric forms [35]. Regarding S100B, Zn^{2+} and Cu^{2+}, which bind to the same site, have quite distinct effects in the amyloid formation by this protein (Figure 10.1): Zn^{2+} has no influence but Cu^{2+} greatly decreases ThT reactivity, implying a structure with lower β-sheet content. It is known that Cu^{2+} drives the oxidation of the two cysteines in the S10B sequence to the disulfide form [36]. This modification is found *in vivo* and is a prerequisite for S100B's neurite extension function [31]. We can hypothesize that the conformational

restriction coupled to disulfide formation partially/completely inhibits the amyloidogenic conversion or induces an alternative conformation. Similar effects may apply to S100A6. Ca^{2+} induced a β-sheet rich aggregated conformation which was not productive for amyloid species and reversed apo-S100A6 cytotoxicity. In the framework of putative S100 cytotoxicity *in vivo*, this finding represents a possibility for the modulation of S100 cytotoxicity by metal ions or other ligands. Apart from metal ions, other relevant modulators of S100 amyloidogenic processes can be interaction with other proteins, some of which occurring even in the absence of calcium and frequently involving the amyloidogenic helices H_I and H_{IV}, or S100 heterodimerization [2]. The reported amyloid formation by S100A8/A9 has been connected with heterodimerization, as the homodimeric counterparts exhibited lower amyloidogenic propensity [7].

In future studies, a systematic and comprehensive description of amyloid formation by S100 proteins will combine the spectroscopic description of amyloidogenesis kinetics (FT-IR, dynamic light scattering, circular dichroism, intrinsic fluorescence) and secondary structure changes (FT-IR and far UV CD spectra deconvolution) to the immunological description of the native-oligomer-fibril transition.

10.6. Conclusions

Here we have described the modulation of amyloidogenesis and amyloid structure of S100A2, S100A3, S100A4, S100A6, S100A12 and S100B by metal ions. All proteins exhibited metal-dependent amyloid endpoint morphologies, even though the formation kinetics is not affected in such a sensitive manner. Amyloid formation by S100A6 was analyzed in more detail, revealing that inhibition of amyloidogenesis by Ca^{2+} was associated to a distinct off-pathway conformation which nevertheless had increased β-sheet content. We quantified the cytotoxicity of native, oligomeric and endpoint amyloid structures of s100A6, S100A12 and S100B, proteins

overexpressed in ALS and Alzheimer's disease [10-13], finding that Ca^{2+} can also revert the cytotoxicity apo-S100A6 oligomers and S100A12 amyloid endpoint structures. From our results, a precise structure-toxicity relationship cannot still be established. This will require future studies which, on the one hand, determine the detailed, molecular level structure of S100 amyloids and, on the other hand, pinpoint the interaction potential of these structures with diverse cell components and the cell signaling pathways being triggered.

It is established that Ca^{2+}, Zn^{2+} and Cu^{2+} promote conformational changes within the S100 fold which have an impact in protein stability (as in S100A2 [19]), in the formation of functional oligomers (as in S100B [5]) and in the formation of amyloid fibrils (as in S100A8/A9 [7]). Considering the latent propensity encoded in the primary sequence of S100 proteins to form β-rich oligomers and fibrils, it is reasonable to envisage that factors like an imbalance in metal homeostasis and anomalous protein-metal interactions, inflammation, oxidative stress or/and genetic mutations may provide conditions in the cellular milieu that affect any of the functional states of S100 proteins (Figure 10.6) and result in the formation of amyloid structures or of its precursor oligomers in a physiological context.

One interesting aspect, which remains to be addressed and may even suggest a toxic gain of function characteristic to amyloid oligomers in general [37], is if S100 amyloids would exacerbate the apoptoptic activity of the S100A8/A9 complex [38-40] or interact with the RAGE receptors further contributing or abrogating the toxic effects. The latter are already known to be involved in Aβ peptide amyloid transport and recognition processes in the context of Alzheimer's disease. A contrasting perspective can also be hypothesized: considering that most of the S100 proteins have upregulated expression patterns in inflammatory, neurodegenerative and malignant proliferation processes, could amyloid formation serve as a sink for

dangerous or somehow harmful proteins promoting inflammation or involved in cancer? Now that even Aβ plaques are viewed from a positive side [41], is it possible that the amyloid formation of S100 proteins may potentially play some "positive" role? Future research in the coming years will certainly contribute to clarify some of these and other questions and will ultimately bring us to higher level of understanding the biology of tumor and degeneration and enable to use our acquired knowledge of S100 structure and functions in developing strategies to modulate their activity for therapeutic purposes.

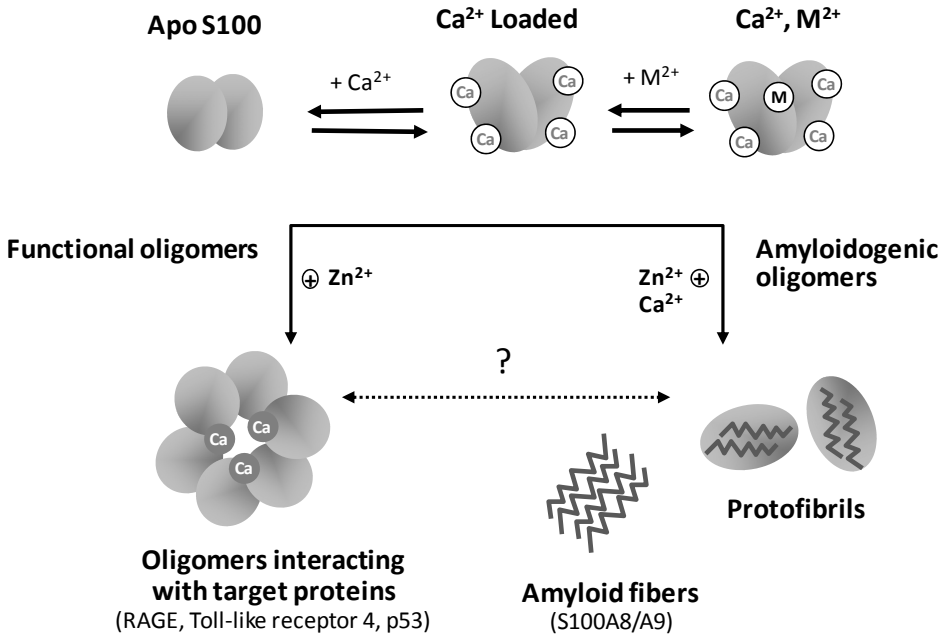


Figure 10.6 - Native states and oligomerization pathways in S100 proteins. Scheme outlining interconversion pathways of S100 proteins, evidencing Ca^{2+} and other metal (M^{2+}) binding sites, and possible routes for oligomerization pathways.

10.7. References

1. Fritz G & Heizmann CW (2004) 3D structures of the calcium and zinc binding S100 proteins. In *Handbook of metalloproteins* (Messerschmidt A, Huber R, Poulos T & Wieghardt K, eds). John Wiley & Sons.
2. Santamaria-Kisiel L, Rintala-Dempsey AC & Shaw GS (2006) Calcium-dependent and -independent interactions of the S100 protein family. *The Biochemical journal* **396**, 201-214

3. Korndörfer IP, Brueckner F & Skerra A (2007) The crystal structure of the human (S100A8/S100A9)2 heterotetramer, calprotectin, illustrates how conformational changes of interacting alpha-helices can determine specific association of two EF-hand proteins. *J Mol Biol* **370**, 887-898
4. Moroz OV, Antson AA, Dodson EJ, Burrell HJ, Grist SJ, Lloyd RM, Maitland NJ, Dodson GG, Wilson KS, Lukanidin E, et al. (2002) The structure of S100A12 in a hexameric form and its proposed role in receptor signalling. *Acta Crystallogr D Biol Crystallogr* **58**, 407-413
5. Ostendorp T, Leclerc E, Galichet A, Koch M, Demling N, Weigle B, Heizmann CW, Kroneck PM & Fritz G (2007) Structural and functional insights into RAGE activation by multimeric S100B. *Embo J* **26**, 3868-3878
6. Koch M, Bhattacharya S, Kehl T, Gimona M, Vasak M, Chazin W, Heizmann CW, Kroneck PM & Fritz G (2007) Implications on zinc binding to S100A2. *Biochim Biophys Acta* **1773**, 457-470
7. Yanamandra K, Alexeyev O, Zamotin V, Srivastava V, Shchukarev A, Brorsson AC, Tartaglia GG, Vogl T, Kayed R, Wingsle G, et al. (2009) Amyloid formation by the pro-inflammatory S100A8/A9 proteins in the ageing prostate. *PLoS One* **4**, e5562
8. Davey GE, Murmann P & Heizmann CW (2001) Intracellular Ca2+ and Zn2+ levels regulate the alternative cell density-dependent secretion of S100B in human glioblastoma cells. *J Biol Chem* **276**, 30819-30826
9. Mandinova A, Atar D, Schafer BW, Spiess M, Aebi U & Heizmann CW (1998) Distinct subcellular localization of calcium binding S100 proteins in human smooth muscle cells and their relocation in response to rises in intracellular calcium. *J Cell Sci* **111 (Pt 14)**, 2043-2054
10. Boom A, Pochet R, Authelet M, Pradier L, Borghgraef P, Van Leuven F, Heizmann CW & Brion JP (2004) Astrocytic calcium/zinc binding protein S100A6 over expression in Alzheimer's disease and in PS1/APP transgenic mice models. *Biochim Biophys Acta* **1742**, 161-168
11. Hoyaux D, Boom A, Van den Bosch L, Belot N, Martin JJ, Heizmann CW, Kiss R & Pochet R (2002) S100A6 overexpression within astrocytes associated with impaired axons from both ALS mouse model and human patients. *J Neuropathol Exp Neurol* **61**, 736-744
12. Mrak RE & Griffinbc WS (2001) The role of activated astrocytes and of the neurotrophic cytokine S100B in the pathogenesis of Alzheimer's disease. *Neurobiol Aging* **22**, 915-922
13. Shepherd CE, Goyette J, Utter V, Rahimi F, Yang Z, Geczy CL & Halliday GM (2006) Inflammatory S100A9 and S100A12 proteins in Alzheimer's disease. *Neurobiol Aging* **27**, 1554-1563
14. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254
15. Morozova-Roche LA, Zurdo J, Spencer A, Noppe W, Receveur V, Archer DB, Joniau M & Dobson CM (2000) Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *J Struct Biol* **130**, 339-351
16. Malisauskas M, Zamotin V, Jass J, Noppe W, Dobson CM & Morozova-Roche LA (2003) Amyloid protofilaments from the calcium-binding protein equine lysozyme: formation of ring and linear structures depends on pH and metal ion concentration. *J Mol Biol* **330**, 879-890
17. Berridge MV, Tan A, S., McCoy KD & Wang R (1996) The Biochemical and Cellular Basis of Cell Proliferation Assays That Use Tetrazolium Salts. *Biochemica* **4**, 14-19
18. Barth A & Zscherp C (2002) What vibrations tell us about proteins. *Q Rev Biophys* **35**, 369-430
19. Botelho HM, Koch M, Fritz G & Gomes CM (2009) Metal ions modulate the folding and stability of the tumor suppressor protein S100A2. *Febs J* **276**, 1776-1786
20. Baglioni S, Casamenti F, Bucciantini M, Luheshi LM, Taddei N, Chiti F, Dobson CM & Stefani M (2006) Prefibrillar amyloid aggregates could be generic toxins in higher organisms. *J Neurosci* **26**, 8160-8167
21. Guignard F, Mauel J & Markert M (1995) Identification and characterization of a novel human neutrophil protein related to the S100 family. *The Biochemical journal* **309 (Pt 2)**, 395-401
22. Kayed R, Head E, Sarsoza F, Saing T, Cotman CW, Necula M, Margol L, Wu J, Breydo L, Thompson JL, et al. (2007) Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol Neurodegener* **2**, 18
23. Haass C & Selkoe DJ (2007) Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat Rev Mol Cell Biol* **8**, 101-112
24. McLean CA, Cherny RA, Fraser FW, Fuller SJ, Smith MJ, Beyreuther K, Bush AI & Masters CL (1999) Soluble pool of Abeta amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. *Ann Neurol* **46**, 860-866

25. Hoyaux D, Decaestecker C, Heizmann CW, Vogl T, Schafer BW, Salmon I, Kiss R & Pochet R (2000) S100 proteins in Corpora amylacea from normal human brain. *Brain Res* **867**, 280-288
26. Strodel B, Lee JW, Whittleston CS & Wales DJ (2010) Transmembrane Structures for Alzheimer's Abeta(1-42) Oligomers. *J Am Chem Soc*
27. Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, Zurdo J, Taddei N, Ramponi G, Dobson CM & Stefani M (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507-511
28. Kirkitadze MD, Bitan G & Teplow DB (2002) Paradigm shifts in Alzheimer's disease and other neurodegenerative disorders: the emerging role of oligomeric assemblies. *J Neurosci Res* **69**, 567-577
29. Walsh DM, Klyubin I, Fadeeva JV, Cullen WK, Anwyl R, Wolfe MS, Rowan MJ & Selkoe DJ (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* **416**, 535-539
30. Demuro A, Parker I & Stutzmann GE (2010) Calcium signaling and amyloid toxicity in Alzheimer disease. *J Biol Chem* **285**, 12463-12468
31. Winningham-Major F, Staecker JL, Barger SW, Coats S & Van Eldik LJ (1989) Neurite extension and neuronal survival activities of recombinant S100 beta proteins that differ in the content and position of cysteine residues. *J Cell Biol* **109**, 3063-3071
32. Huttunen HJ, Kuja-Panula J, Sorci G, Agneletti AL, Donato R & Rauvala H (2000) Coregulation of neurite outgrowth and cell survival by amphoterin and S100 proteins through receptor for advanced glycation end products (RAGE) activation. *J Biol Chem* **275**, 40096-40105
33. Kogel D, Peters M, Konig HG, Hashemi SM, Bui NT, Arolt V, Rothermundt M & Prehn JH (2004) S100B potently activates p65/c-Rel transcriptional complexes in hippocampal neurons: Clinical implications for the role of S100B in excitotoxic brain injury. *Neuroscience* **127**, 913-920
34. Sciarretta KL, Gordon DJ, Petkova AT, Tycko R & Meredith SC (2005) Abeta40-Lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid. *Biochemistry* **44**, 6003-6014
35. Noy D, Solomonov I, Sinkevich O, Arad T, Kjaer K & Sagi I (2008) Zinc-amyloid beta interactions on a millisecond time-scale stabilize non-fibrillar Alzheimer-related species. *J Am Chem Soc* **130**, 1376-1383
36. Matsui Lee IS, Suzuki M, Hayashi N, Hu J, Van Eldik LJ, Titani K & Nishikimi M (2000) Copper-dependent formation of disulfide-linked dimer of S100B protein. *Arch Biochem Biophys* **374**, 137-141
37. Stefani M (2010) Amyloid cytotoxicity: contributions of both oligomer/fibril and cell membrane biochemical and biophysical features. *FEBS J*, mini-review series
38. Ghavami S, Eshragi M, Ande SR, Chazin WJ, Klonisch T, Halayko AJ, McNeill KD, Hashemi M, Kerkhoff C & Los M (2010) S100A8/A9 induces autophagy and apoptosis via ROS-mediated cross-talk between mitochondria and lysosomes that involves BNIP3. *Cell Res* **20**, 314-331
39. Ghavami S, Kerkhoff C, Chazin WJ, Kadkhoda K, Xiao W, Zuse A, Hashemi M, Eshraghi M, Schulze-Osthoff K, Klonisch T, et al. (2008) S100A8/9 induces cell death via a novel, RAGE-independent pathway that involves selective release of Smac/DIABLO and Omi/HtrA2. *Biochim Biophys Acta* **1783**, 297-311
40. Viemann D, Barczyk K, Vogl T, Fischer U, Sunderkotter C, Schulze-Osthoff K & Roth J (2007) MRP8/MRP14 impairs endothelial integrity and induces a caspase-dependent and -independent cell death program. *Blood* **109**, 2453-2460
41. Heo C, Chang KA, Choi HS, Kim HS, Kim S, Liew H, Kim JA, Yu E, Ma J & Suh YH (2007) Effects of the monomeric, oligomeric, and fibrillar Abeta42 peptides on the proliferation and differentiation of adult neural stem cells from subventricular zone. *J Neurochem* **102**, 493-500

11. Cross-talk between S100B and Aβ amyloidogenesis

Proteins were purified by Gunter Fritz (Freiburg University, DE). AFM analysis was performed by Kiran Yanamandra and Ludmilla A. Morozova-Roche (Umeå University, SE).

11.1. Summary

The recently identified amyloidogenic properties of S100 proteins related to *corpora amylacea* formation combined with the presence of numerous S100 proteins in *corpora amylacea* in the brain raises the possibility of the involvement of other S100 proteins in pathological processes. In particular, the association of S100B with Alzheimer's disease and accumulation in the synaptic cleft support the hypothesis that amyloid cross-seeding between oligomers of these proteins may be relevant in disease. Here we carry out S100B and $A\beta_{1-40}$ cross-seeding experiments to address this possibility. Our results show that $A\beta_{1-40}$ fibrils – but not oligomers or other amyloid fibrils – seed S100B amyloidogenesis. Conversely, S100B amyloids specifically seed $A\beta_{1-40}$ fibrillization. Overall, our data supports the hypothesis of correlation between Aβ and S100B amyloid formation including seeding effects facilitating cytotoxicity onset and neurodegeneration.

11.2. Introduction

Protein deposition as amyloid oligomers in the brain is the cytological hallmark of several neurodegenerative diseases including Alzheimer's (AD), Parkinson's (PD) and Amyotrophic Lateral Sclerosis (ALS). The identification of cellular modulators of protein deposition, either proteinaceous or not, remains a challenging issue. In this respect, the recently identified amyloid forming properties of S100 proteins opens new possibilities in regard to the involvement of these proteins in neurodegenerative processes [1]. Due to their cell- and tissue-specific expression patterns, S100 proteins are frequently used as disease markers. This includes the above mentioned neurodegenerative conditions: S100A6 [2], S100A12 [3] and S100B [4] are overexpressed in the brains of Alzheimer's disease patients; S100A6 is found overexpressed in astrocytes located near impaired axons of motor neurons

of ALS patients [5]; and the serum levels of S100B correlate with PD severity [6-7].

S100B is highly abundant in the brain, constituting about 0.5% of all brain protein [8]. It is abundantly expressed in astrocytes and released constitutively [9], although other cell types also express S100B [10]. Schwann cells release S100B following RAGE activation [11-12]. Astrocytes release S100B via activation of metabotropic glutamate receptor 3 in a neural- and synaptic-activity-dependent manner [13]. Numerous modulators of astrocytic S100B secretion are known. Stimulation by Ca^{2+} [14], antioxidants [15], branched-chain α-keto acids [16], 5-HT_{1A} [17], lysophosphatidic acid [18], glutamate [19] and TNF-α [20], and during metabolic stress [21] enhance S100B release. High glucose [22] or glutamate [23] concentrations, inhibition of Src kinase [24] or cell confluency [9, 14] inhibit S100B secretion. Extracellular S100B concentration determines protein function *in vivo*. At the physiological nanomolar level, S100B is a neurotrophic agent promoting neuronal survival and differentiation [25]. Micromolar concentrations are pro-apoptotic [26].

Enhanced release of S100B into the brain extracellular space is associated with inflammatory foci including traumatic cortical injury or focal cerebral ischemia, some psychiatric disorders [27] or Down's syndrome [28], the latter being due to the extra copy of the S100B gene in chromosome 21. The connection of S100B with Alzheimer's disease involves the two main proteins implicated in the pathology: Aβ and Tau. These two interactors represent the two cellular locations for S100B: intracellular (Tau) and extracellular (Aβ). The formation of extracellular Aβ amyloid deposits induces astrogliosis – a general response of the central nervous system (CNS) to neuronal injury [29-30] which consists in the enhancement of the size and number of GFAP-expressing astrocytes (*i. e.* reactive astrocytes) – around Aβ amyloid deposits [4] and local accumulation of S100B. S100B increases

amyloid precursor protein (APP) expression [31], it is implicated in the formation of dystrophic neurites in Aβ amyloid plaques [32] and exacerbates Aβ neurotoxicity [33]. Overexpression of S100B alone accelerates the onset of AD-like pathology in transgenic mouse models [34] through increasing the levels of the Aβ peptide and of brain parenchymal and cerebral vascular Aβ deposits. On the other hand, Aβ stimulates S100B synthesis [35].

The recent discovery of amyloid formation by S100A8/A9 within *corpora amylacea* in the prostate [1] established the proof of principle for the correlation of such structures and S100 amyloidogenesis. Notably, the healthy brain contains *corpora amylacea* [36] which contain numerous S100 proteins: S100A1, S100A2, S100A3, S100A4, S100A5, S100A6, S100A8, S100A9 and S100A12 as well as APP [37]. The structure of these proteins in *corpora amylacea* is currently unknown. S100B is not detected in *corpora amylacea* but is found in the nearby tissue due to astrocytosis. *Corpora amylacea* become more abundant in Alzheimer's [38] and ALS [39] patients' brains. Their functions and what triggers their appearance is not clear. All these evidences strongly suggest that S100B is directly involved in the process of amyloid formation by Aβ. In fact, S100 proteins have intrinsic amyloidogenic potentials similar to Aβ peptides [40] and S100A8/A9 have been found to form amyloid *in vivo* [1].

Here we report groundbreaking data regarding amyloid formation by the S100 proteins which are overexpressed in the brain in AD and ALS: S100A6, S100A12 and S100B [2-5]. Building on our findings that these S100 proteins form amyloid structures, we correlate this behavior with the amyloidogenic properties of the $A\beta_{1-40}$ peptide, forming the senile plaques in AD. Preformed $A\beta_{1-40}$ fibrils seed amyloid formation by S100B. More interestingly, $A\beta_{1-40}$ fibrillization is enhanced specifically by preformed S100B fibrils in physiological-like conditions, as lysozyme fibrils had an effect similar to an

unseeded control. The possible implications of the involvement of S100B in AD through modulating Aβ fibrillization are discussed.

11.3. Materials and methods

11.3.1. Chemicals and proteins

All reagents were of the highest grade commercially available. Thioflavin T was obtained from Sigma. Trace metals were removed from all solutions by solution treatment with a chelex resin. S100B was expressed in *E. coli* and purified to homogeneity as described previously [41]. S100B molar concentrations refer to the subunit. Hen egg white lysozyme was obtained from Fluka. $A\beta_{1-40}$ was obtained from Alexotech (Umeå, Sweden). 0.5 mg were carefully dissolved in 10 mM NaOH pH 7.5, quickly frozen at liquid nitrogen temperature and stored at -80°C until usage.

11.3.2. Protein fibrillization

Protein fibrillization was promoted at different conditions, dependent on the intrinsic amyloidogenic potential of each protein. S100B solutions were prepared in 50 mM glycine pH 2.5 and set at 3 mg/ml using Bradford's method [42]. Amyloidogenesis was promoted by quiescent incubation at 42°C, 52°C or 57°C. $A\beta_{1-40}$ fibrillization was promoted by incubation at 25°C in 10 mM NaOH pH 7.5. In either case, samples were kept at 4°C at all times before amyloid formation assays. Two cross-seeding experiments were carried out: (i) incubating S100B in 50 mM glycine pH 2.5 at 42°C in the presence of 5% (w/w) preformed $A\beta_{1-40}$ fibrils; (ii) incubating $A\beta_{1-40}$ in 10 mM NaOH pH 7.5 at 25°C in the presence of 5% (w/w) preformed fibrils of S100B or lysozyme [43]. Before each assay, any pre-existing aggregates were pelleted and removed by centrifuging at 12000 g for 10 minutes. Lysozyme amyloid seeds were obtained by incubating 10 mg/ml protein in 50 mM glycine pH 2.5 at 57°C for two weeks [43]. $A\beta_{1-40}$ amyloid seeds were

obtained by incubating 2 mg/ml $A\beta_{1-40}$ in 10 mM NaOH pH 7.5 at 25°C for 3 days.

11.3.3. Thioflavin T fluorescence

ThT binding assays were performed according to [43]. Briefly, 5 µl of protein sample (15 µg) was mixed with 65 µM ThT in 10 mM potassium phosphate (KPi), 150 mM NaCl, pH 7.0 at room temperature to a final volume of 1 ml. Samples were incubated under stirring for 1-2 min and ThT fluorescence at 482 nm was recorded with a Cary Varian Eclipse instrument with temperature set at 25°C using a peltier-thermostated cell support. Excitation wavelength was 440 nm, PMT voltage 600 V, and excitation and emission slits 10 nm.

11.3.4. Atomic force microscopy

After reaching the plateau phase in amyloid formation kinetics, samples were stored at 4°C. Before analysis, S100B (3 mg/ml) samples were diluted 100 fold and $A\beta_{1-40}$ (2 mg/ml) samples were diluted 50 fold in water and deposited on a mica surface. Morphology was analyzed by AFM on a PICO PLUS 5500 microscope (Agilent) as previously described [44].

11.3.5. ATR FT-IR spectroscopy

Secondary structural changes occurring during amyloid formation were monitored by ATR FT-IR measurements performed in a Bruker IFS 66/S spectrometer equipped with a nitrogen-cooled MCT detector using the thermostatized Harrick BioATR cell II. 3 mg/ml S100 proteins in 50 mM glycine pH 2.5 were incubated at 57°C while spectra were continuously acquired (2 min accumulation time, 12 mm aperture, 20 kHz scanner velocity). Spectral resolution was 4 cm^{-1}. Amyloid formation was assessed by the absorption variation at β-sheet associated wavenumbers (1628 cm^{-1} for S100B, 1627 cm^{-1} for $A\beta_{1-40}$) normalized by the absorption at the amide I spectral maximum (1653 cm^{-1} for S100B, 1649 cm^{-1} for $A\beta_{1-40}$). To identify

the spectral components changing over time, spectra where deconvoluted using singular value decomposition implemented on Matlab [45].

11.4. Results

11.4.1. Lower temperatures slow S100B amyloidogenesis without changing the morphology of oligomers

To explore the hypothesis that the amyloidogenic processes of S100B and $A\beta_{1-40}$ may be correlated, we have devised an *in vitro* experimental approach combining amyloid formation assays detected by thioflavin T (ThT) fluorescence, FT-IR and atomic force microscopy (AFM). This approach is based on cross-seeding assays, a methodology currently employed to detect proteins interacting with Aβ and regulating fibrillization [46]. More importantly, seeding is a fundamental characteristic of amyloidogenic processes [47] and allows analyzing the mechanism of amyloid formation. Our hypothesis is that spontaneous formation of amyloid species by S100B or Aβ – which are both overexpressed in the synaptic milieu in Alzheimer's disease – may cross-seed amyloid formation, a process with potential implications in disease progression.

Amyloid seeding is a methodology wherein the lag phase preceding amyloid formation is shortened by the addition of pre-formed amyloid nuclei. We have previously determined that S100B rapidly forms oligomer or protofibrilllar amyloid species at 57°C and pH 2.5 (Chapters 9, 10). This limits the sensitivity of cross-seeding assays, which rely on the possibility of the amyloidogenesis kinetics being accelerated. Since amyloid formation is dependent on temperature, we have analyzed amyloid formation by S100B at lower temperatures than the one previously selected (Figure 11.1). By lowering the temperature from 57 to 42°C the kinetics of amyloid formation by S100B is significantly slowed down without important differences in the steady state ThT fluorescence intensity. More importantly, structural

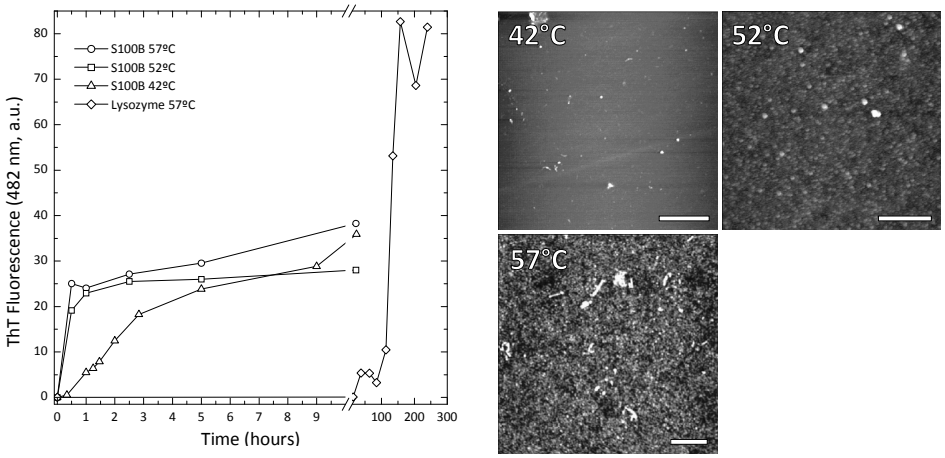


Figure 11.1- Temperature dependence of S100B amyloidogenesis at pH 2.5. (Left) Kinetic traces from ThT fluorescence assays. S100B was 3 mg/ml in 50 mM glycine pH 2.5. Lysozyme 10 mg/ml in 50 mM glycine pH 2.5 is the positive control for amyloid formation. (Right) AFM images of S100B amyloid samples after incubation at the respective temperature. Scale bar: 1 μm.

analysis of S100B samples at the end stage amyloid state revealed identical morphologies of the oligomers formed by S100B, independently of incubation temperature.

11.4.2. $A\beta_{1-40}$ fibrils seed S100B amyloidogenesis

Using the conditions previously optimized for S100B amyloidogenesis (3 mg/ml S100B, 50 mM glycine pH 2.5), we have assessed the seeding effects of $A\beta_{1-40}$. Taking into account the structural diversity of Aβ amyloid structures [48], we have seeded S100B with oligomeric and fibrillar $A\beta_{1-40}$. Lysozyme fibrils were used in a control assay. By monitoring S100B amyloidogenesis using ThT fluorescence, we observed that $A\beta_{1-40}$ fibrils (5% w/w) specifically accelerated amyloidogenesis onset (Figure 11.2A), unlike oligomeric $A\beta_{1-40}$ or lysozyme fibrils. Examination of the amyloid morphology after the seeding assay revealed that the oligomeric structure of S100B was kept irrespectively of the presence of seeds (Figure 11.3). This is direct evidence that amyloidogenesis seeding by $A\beta_{1-40}$ is a purely kinetic effect and not an artifact arising from a distinct final structure in seeded S100B.

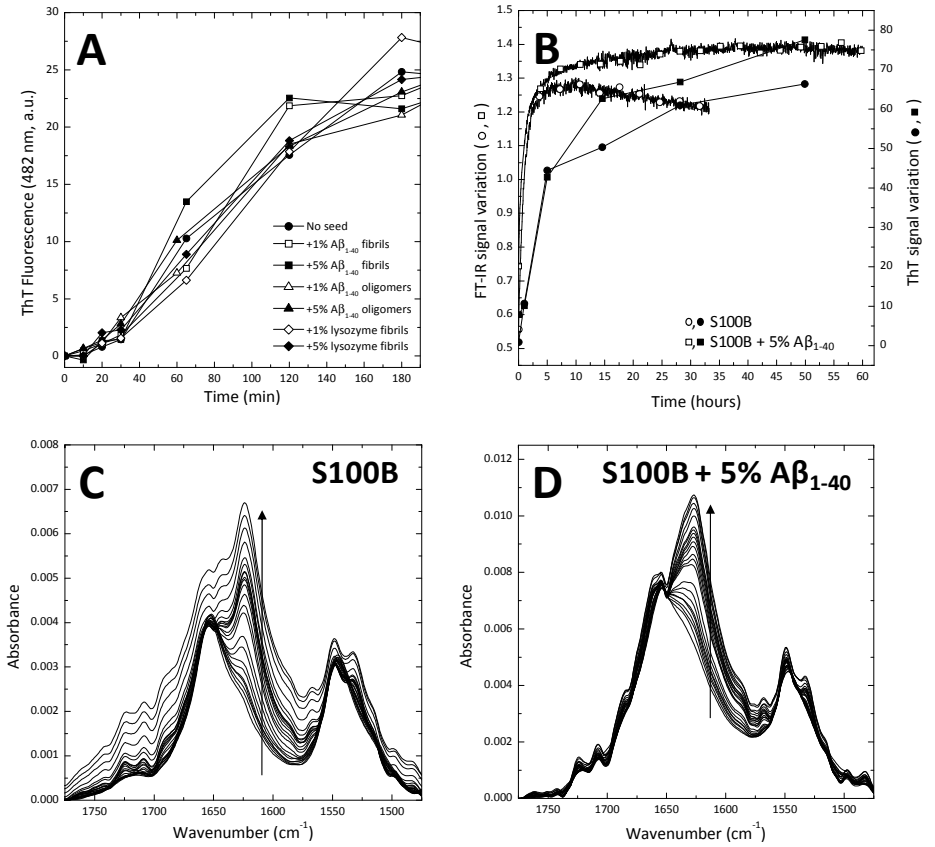


Figure 11.2 - Seeding of S100B amyloidogenesis with $A\beta_{1\text{-}40}$. S100B 3 mg/ml in 50 mM glycine pH 2.5 was incubated at 42°C alone or in the presence of 1 or 5% (w/w) preformed $A\beta_{1\text{-}40}$ fibrils, $A\beta_{1\text{-}40}$ amyloid oligomers or lysozyme fibrils. Amyloidogenesis kinetics assessed by ThT fluorescence (A) or ThT fluorescence and FT-IR signal variation at 1628 cm^{-1}, normalized by the absorption at 1653 cm^{-1} (B). Time-resolved FT-IR spectra of unseeded S100B (C) and S100B seeded with 5% $A\beta_{1\text{-}40}$ fibrils (D).

Nevertheless, the seeding effect is modest. To gather information about the secondary structure changes occurring in S100B during amyloidogenesis, we have examined the amyloidogenesis of S100B seeded with 5% $A\beta_{1\text{-}40}$. In seeded and unseeded S100B, amyloid formation can be monitored by the increasing absorbance at 1628 cm^{-1} (Figure 11.2CD), arising by the buildup of intra and intermolecular β-sheet content. When amyloidogenesis assays were allowed to progress until a plateau phase was reached, higher ThT and FT-IR signals were recorded in the seeded sample (Figure 11.2B), which is an indication of enhanced β-sheet content. Since the endpoint morphology is not

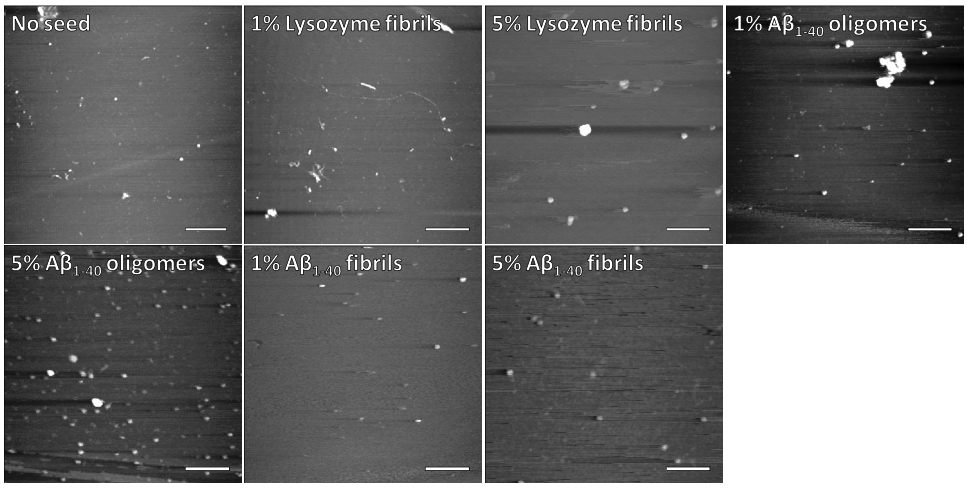


Figure 11.3 - Morphology of seeded S100B amyloid oligomers. From AFM imaging. S100B was prepared at 3 mg/ml in 50 mM glycine pH 2.5 and incubated at 42°C in the absence and presence of 1 or 5% (w/w) preformed lysozyme fibrils, $A\beta_{1\text{-}40}$ amyloid oligomers or $A\beta_{1\text{-}40}$ amyloid fibrils. Images were taken at the plateau phase of amyloid formation kinetics. Scale bar: 1 μm.

affected (Figure 11.3), this result can be interpreted as an indication of enhanced amyloid oligomerization efficiency.

We have obtained further information from FT-IR spectra using singular value decomposition (SVD) deconvolution [45]. This algebraic algorithm rearranges the succession of FT-IR spectra we have recorded during S100B amyloidogenesis into the temporal variation of its components. The relative intensity of each component indicates which ones contain data representative of the experimental result and which ones represent noise. From our analysis, only the first three SVD components of the unseeded and seeded S100B amyloidogenesis assays contained significant information (Figure 11.4AB). Using SVD we could obtain additional structural information. The first component is representative of the final amyloid state. The second component of both experiments contains a peak at a wavenumber associated with intermolecular β-sheets (1617 cm^{-1} unseeded, 1623 cm^{-1} seeded), which are characteristic of protein aggregates and particularly amyloid [49]. The kinetic analysis of this component revealed

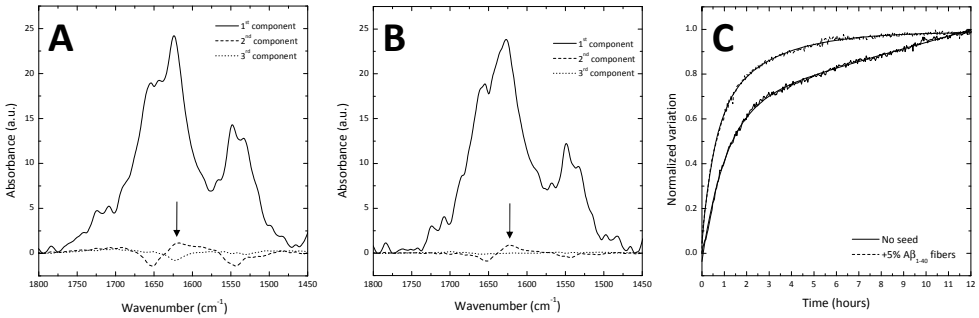


Figure 11.4 - Singular value analysis of FT-IR monitored S100B amyloidogenesis seeding by Aβ$_{1\text{-}40}$. The spectra acquired during S100B amyloidogenesis (Figure 11.2CD) were deconvoluted using SVD and the spectral components changing with time were identified. Main spectral components of unseeded S100B (A) and S100B seeded with 5% Aβ$_{1\text{-}40}$ fibrils (B). (C) Normalized variation of component 2, which contains a significant contribution at the β-sheet associated peak around 1620 cm^{-1} (arrows).

that its intensity increased with time and that buildup of intermolecular β-sheets is faster when fibrillar Aβ$_{1\text{-}40}$ seeds are present (Figure 11.4C).

11.4.3. S100B is a specific seed of Aβ$_{1\text{-}40}$ fibrillization in vitro

The seeding of S100B amyloidogenesis with Aβ$_{1\text{-}40}$ fibrils is an indication of the interaction between these two proteins. In the case of amyloid formation, this interaction can be assumed to be based on the existence of a specific template provided by the seed. Accordingly, a comprehensive analysis of the interaction potential of S100B and Aβ$_{1\text{-}40}$ requires the assessment of the converse seeding experiment: seeding Aβ$_{1\text{-}40}$ amyloidogenesis with preformed S100B fibrils.

Following the same approach, we have monitored Aβ$_{1\text{-}40}$ fibrillization at optimized conditions (2 mg/ml, 10 mM NaOH pH 7.5, 25°C) using ThT fluorescence and FT-IR. After fibrillization, samples were imaged using AFM. Aβ$_{1\text{-}40}$ was seeded with 5% (w/w) preformed fibrils of S100B or lysozyme. Like S100B, Aβ$_{1\text{-}40}$ fibrillization results in the development of an FT-IR absorbance band at 1627 cm^{-1} (Figure 11.5A-C), typical of intra and intermolecular β-sheets. AFM analysis of the wither sample after amyloid seeding assays did not allow an unequivocal structural characterization

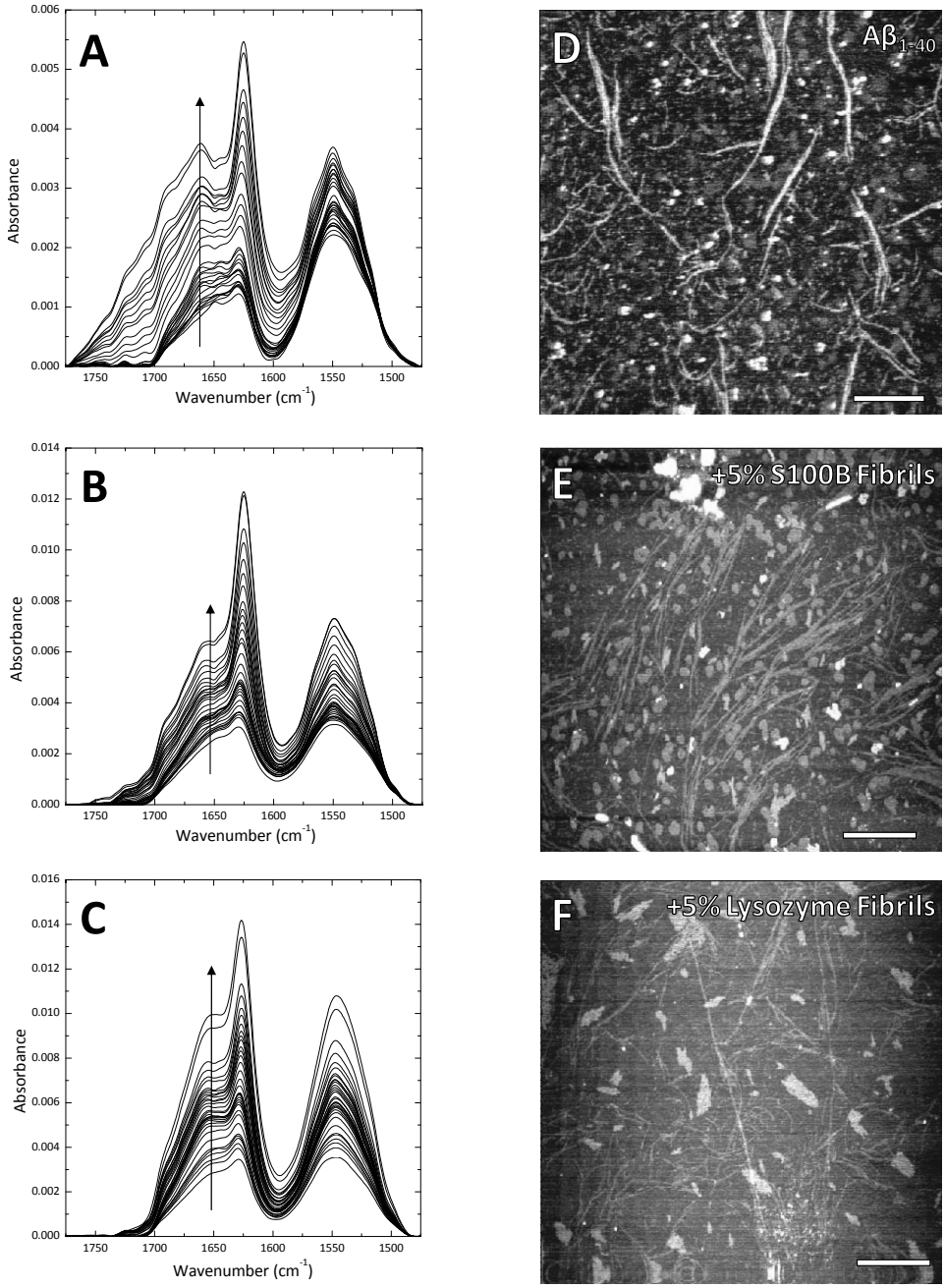


Figure 11.5 - Aβ$_{1-40}$ amyloidogenesis seeding. 5% (w/w) fibrils of S100B or lysozyme were used as seeds. (A-C) FT-IR spectra recorded during amyloid formation. (D-F) AFM images obtained after the seeding assays. (A,D) Unseeded Aβ$_{1-40}$; (B,E) Aβ$_{1-40}$ seeded with 5% S100B fibrils; (C,F) Aβ$_{1-40}$ seeded with 5% lysozyme fibrils.

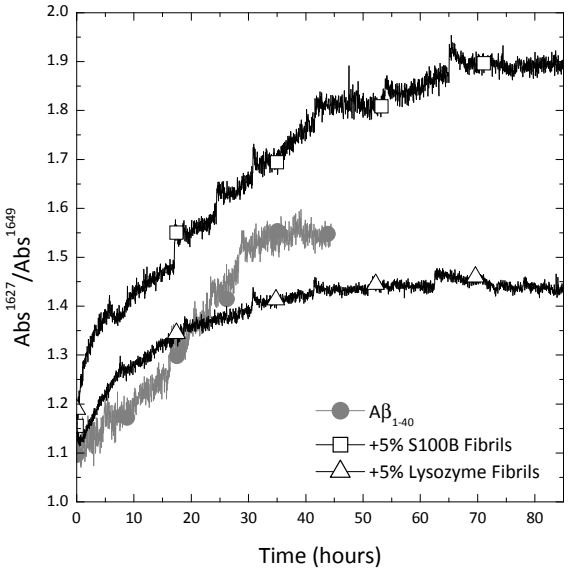


Figure 11.6 - $A\beta_{1-40}$ amyloid seeding kinetics. $A\beta_{1-40}$ (2mg/ml, 10 mM NaOH pH 7.5) was incubated at 25°C with and without 5% (w/w) fibrillar S100B or lysozyme. Fibrillization kinetics was assessed from the absorbance change at 1627 cm^{-1} normalized by the absorbance at 1649 cm^{-1}.

because of sample heterogeneity, with fibrillar and aggregate structures around 1 nm in height in all samples. AFM is not a quantitative technique, as the detected proteins are those which bind the negatively charged mica surface, possibly not accurately representing the whole solution ensemble. This excludes the possibility of a quantitative appreciation of the images. Heterogeneity also precludes the immediate identification of the identity of the proteins composting the fibrillar and the aggregate structures.

Significant seeding effects occur in the amyloid formation kinetics (Figure 11.6): S100B seeds both accelerate $A\beta_{1-40}$ fibrillization and increase the β-sheet content of the sample ensemble, as judged by the higher signal at the plateau phase. More importantly, comparing with the results obtained using lysozyme seeds, the effect is S100B-specific.

11.5. Discussion

The results reported herein show that S100B and $A\beta_{1-40}$ amyloids can cross-seed the amyloid formation by the other protein, establishing the proof of concept for the possible functional implications of the interaction of the two proteins *in vivo*. We have based our analysis of S100B-Aβ interaction on cross-seeding assays because seeding requires specific structural templating.

Such approach has already been pursued in the literature for identifying Aβ interacting partners. Stefin B was recently shown to modulate Aβ fibrillization [46]. Substoichiometric amounts of stefin B dimers, tetramers and high order oligomers inhibit Aβ fibrillization, while other oligomeric forms are not. Similar effects exist for $A\beta_{1-40}$ and $A\beta_{1-42}$ cross-seeding. Monomeric $A\beta_{1-40}$ and $A\beta_{1-42}$ can be seeded by fibrils of either peptide. However, $A\beta_{1-42}$ protofibrils selectively seed amyloid formation by monomeric $A\beta_{1-42}$ but not monomeric $A\beta_{1-40}$. On the other hand, monomeric $A\beta_{1-40}$ alters the kinetic stability, solubility, and morphology of $A\beta_{1-42}$ aggregates and prevents fibril assembly [50]. Accordingly, the ratio of $A\beta_{1-40}$ to $A\beta_{1-42}$ influences the amyloidogenic and, consequently, toxic properties of Aβ. This demonstrates that seeding effects are not related to protein identity but to stringent structural determinants.

Apart from the extensive data correlating S100B and Aβ in Alzheimer's disease, reports also exist on the seeding effects of Aβ. Murine cortical neurons and human neuroblastoma cells (SHSY-5Y) can uptake $A\beta_{1-42}$ (≥1 nM) from the extracellular medium and concentrate it in late endosomes or lysosomes to micromolar levels, where aggregates with monomeric $A\beta_{1-42}$ seeding capacity assemble [51]. Such intracellular accumulation of $A\beta_{1-42}$ in AD-sensitive brain areas precedes disease onset, suggesting an early role in AD neurodegeneration [52].

The importance of the regulation of fibrillization lag phase has been highlighted by the work of Eaton and co-workers in sickle cell anemia in the 1970s. The aggregation of mutant hemoglobin which underlies disease is formally dependent on approximately the 30^{th} power of the soluble precursor concentration [53], implying that even small variations in protein concentration or temperature have very significant effects in the aggregation kinetics. The still unaddressed kinetic description of S100B and $A\beta_{1-40}$ cross seeding thus encloses important information about the relevance of such

effects. Of particular interest is the putative deregulation of S100B signaling by Aβ-related amyloid seeding. RAGE binds Aβ [54], S100B [55] and other S100 proteins [56]. In the case of Aβ, the RAGE interaction is involved in the translocation of circulating Aβ to the brain [57]. Also, RAGE binding potentiates Aβ-induced neuronal dysfunction [58]. Nanomolar S100B concentrations protect neurons against Aβ toxicity in a RAGE-dependent manner [59], involving the overexpression of anti-apoptotic protein Bcl2 and, possibly, Aβ scavenging by enhanced soluble RAGE (sRAGE) secretion [60]. However, Aβ senile plaques initiate an inflammatory response leading to the local overexpression of proteins like S100A6, S100A8/A9, S100A12 and S100B [56]. Micromolar S100B concentrations are characteristic of this condition and present additional toxicity relatively to Aβ alone [59].

In the same way that overexpression and extracellular co-localization of amyloidogenic S100B and Aβ in Alzheimer's disease suggested us that the amyloidogenic processes of both could be correlated, similar inferences can be drawn for other S100 proteins with pathological relevance. In ALS, another neurodegenerative disease where intraneuronal Cu/Zn Superoxide Dismutase (SOD1) aggregation is associated with motor neuron loss [61]. In mice [36] and humans [5], ALS is accompanied by astrogliosis and overexpression of S100A6 in astrocytes. Like the normal population, ALS patients have c*orpora amylacea* in the spinal cord [62-63] and brain stem [64-65], but present unique deposits in intramuscular nerves [63, 66], the composition of which is currently unknown.

Alzheimer's disease is characterized by an imbalancement in metal ion homeostasis. Specifically, zinc transporters co-localize with amyloid plaques of Alzheimer's disease brain [67] and zinc induces Aβ amyloid formation. Like Aβ, S100B is also a zinc-binding protein. The other S100B metal ion ligands – calcium and copper – are abundant in the synaptic space where Aβ deposits form in AD. Similarly to the strictly metal-dependent amyloid

formation by S100A8/A9, metal ions may be relevant in this protein's amyloid processes.

Structural characterization of S100B and Aβ cross-seeded species was incomplete in this work. The elucidation of the structure of the seeded amyloid samples was hampered by structural heterogeneity because of the multiple species in solution. A detailed characterization of S100B and $A\beta_{1-40}$ cross-seeded amyloids will thus require characterization and identification of individual oligomeric and fibrillar species. Finally, the validation of our hypothesis on the cross-talk between S100 and Aβ amyloidogenesis in Alzheimer's disease will require the investigation of the formation of S100B amyloids at physiological or physiological-like conditions and the examination of AD brain samples to determine if S100B deposition is observed in patients.

11.6. Conclusions

S100B has been recurrently implicated in Alzheimer's disease. Our results suggest that a possible connection could consist in the direct involvement of this protein in Aβ fibrillization. By using a series of amyloid cross-seeding assays we have showed that $A\beta_{1-40}$ fibrils speed amyloid formation by S100B and that fibrils of S100B – but not of the unrelated protein lysozyme – were able to eliminate the lag phase preceding amyloid formation by $A\beta_{1-40}$. These observations support the hypothesis that spontaneous formation of amyloid species by either S100B or Aβ may establish a cross-seeding catalytical cycle promoting amyloid formation and disease progression. This may also apply with S100A6, S100A12 and S100A8/A9, which are overexpressed in Alzheimer's disease as a consequence of the astroglial inflammatory response to the deposition of Aβ plaques. The currently proposed hypothesis will be tested in the future with the determination of the conditions where S100B is prone to amyloid formation, with the search of S100B/Aβ co-deposition *in vivo* and with the identification of modulators of S100B

amyloidogenesis. Regarding the latter, zinc, which is an S100B ligand and promotes Aβ fibrillization is an *a priori* candidate (Figure 11.7).

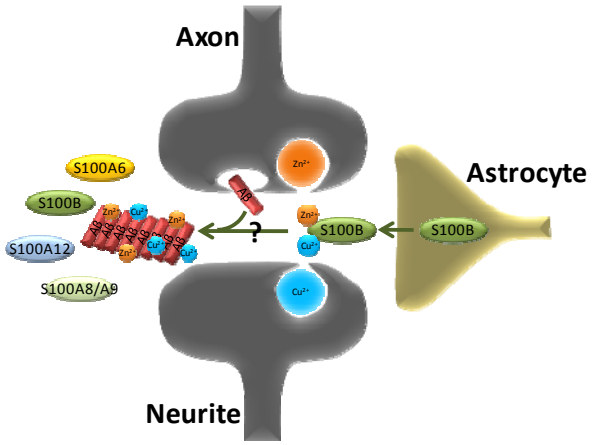


Figure 11.7 - S100 proteins in Alzheimer's disease. Alzheimer's disease is characterized by abundant deposition of extracellular plaques in the central nervous system. These recruit glial cells which secrete S100 proteins to the same extracellular space [56]. Zinc and copper, which are released as μM transients in the glutamatergic synapse [68], enhance fibrillation Aβ fibrillization [69] and bind to S100B. Co-localization of high concentrations of amyloidogenic Aβ and S100 proteins supports the hypothesis that the amyloidogenic processes of these proteins may be related. Thus, formation of Aβ or S100 amyloids may cross-template further fibrillization and favor disease progression.

11.7. References

1. Yanamandra K, Alexeyev O, Zamotin V, Srivastava V, Shchukarev A, Brorsson AC, Tartaglia GG, Vogl T, Kayed R, Wingsle G, et al. (2009) Amyloid formation by the pro-inflammatory S100A8/A9 proteins in the ageing prostate. *PLoS One* **4**, e5562
2. Boom A, Pochet R, Authelet M, Pradier L, Borghgraef P, Van Leuven F, Heizmann CW & Brion JP (2004) Astrocytic calcium/zinc binding protein S100A6 over expression in Alzheimer's disease and in PS1/APP transgenic mice models. *Biochim Biophys Acta* **1742**, 161-168
3. Shepherd CE, Goyette J, Utter V, Rahimi F, Yang Z, Geczy CL & Halliday GM (2006) Inflammatory S100A9 and S100A12 proteins in Alzheimer's disease. *Neurobiol Aging* **27**, 1554-1563
4. Mrak RE & Griffinbc WS (2001) The role of activated astrocytes and of the neurotrophic cytokine S100B in the pathogenesis of Alzheimer's disease. *Neurobiol Aging* **22**, 915-922
5. Hoyaux D, Boom A, Van den Bosch L, Belot N, Martin JJ, Heizmann CW, Kiss R & Pochet R (2002) S100A6 overexpression within astrocytes associated with impaired axons from both ALS mouse model and human patients. *J Neuropathol Exp Neurol* **61**, 736-744
6. Schaf DV, Tort AB, Fricke D, Schestatsky P, Portela LV, Souza DO & Rieder CR (2005) S100B and NSE serum levels in patients with Parkinson's disease. *Parkinsonism Relat Disord* **11**, 39-43
7. Wilhelm KR, Yanamandra K, Gruden MA, Zamotin V, Malisauskas M, Casaite V, Darinskas A, Forsgren L & Morozova-Roche LA (2007) Immune reactivity towards insulin, its amyloid and protein S100B in blood sera of Parkinson's disease patients. *Eur J Neurol* **14**, 327-334
8. Kimura S, Kato K, Semba R & Isobe T (1984) Regional distribution of S-100a(0) (??), S-100a (??) and S-100b (??) in the bovine central nervous tissue determined with a sensitive enzyme immunoassay system. *Neurochem Int* **6**, 513-518

9. Van Eldik LJ & Zimmer DB (1987) Secretion of S-100 from rat C6 glioma cells. *Brain Res* **436**, 367-370
10. Donato R, Sorci G, Riuzzi F, Arcuri C, Bianchi R, Brozzi F, Tubaro C & Giambanco I (2009) S100B's double life: intracellular regulator and extracellular signal. *Biochim Biophys Acta* **1793**, 1008-1022
11. Bierhaus A, Humpert PM, Morcos M, Wendt T, Chavakis T, Arnold B, Stern DM & Nawroth PP (2005) Understanding RAGE, the receptor for advanced glycation end products. *J Mol Med* **83**, 876-886
12. Schmidt AM, Yan SD, Yan SF & Stern DM (2001) The multiligand receptor RAGE as a progression factor amplifying immune and inflammatory responses. *J Clin Invest* **108**, 949-955
13. Sakatani S, Seto-Ohshima A, Shinohara Y, Yamamoto Y, Yamamoto H, Itohara S & Hirase H (2008) Neural-activity-dependent release of S100B from astrocytes enhances kainate-induced gamma oscillations in vivo. *J Neurosci* **28**, 10928-10936
14. Davey GE, Murmann P & Heizmann CW (2001) Intracellular Ca2+ and Zn2+ levels regulate the alternative cell density-dependent secretion of S100B in human glioblastoma cells. *J Biol Chem* **276**, 30819-30826
15. de Almeida LM, Pineiro CC, Leite MC, Brolese G, Tramontina F, Feoli AM, Gottfried C & Goncalves CA (2007) Resveratrol increases glutamate uptake, glutathione content, and S100B secretion in cortical astrocyte cultures. *Cell Mol Neurobiol* **27**, 661-668
16. Funchal C, Tramontina F, Quincozes dos Santos A, Fraga de Souza D, Goncalves CA, Pessoa-Pureur R & Wajner M (2007) Effect of the branched-chain alpha-keto acids accumulating in maple syrup urine disease on S100B release from glial cells. *J Neurol Sci* **260**, 87-94
17. Whitaker-Azmitia PM, Murphy R & Azmitia EC (1990) Stimulation of astroglial 5-HT1A receptors releases the serotonergic growth factor, protein S-100, and alters astroglial morphology. *Brain Res* **528**, 155-158
18. Pinto SS, Gottfried C, Mendez A, Goncalves D, Karl J, Goncalves CA, Wofchuk S & Rodnight R (2000) Immunocontent and secretion of S100B in astrocyte cultures from different brain regions in relation to morphology. *FEBS Lett* **486**, 203-207
19. Ciccarelli R, Di Iorio P, Bruno V, Battaglia G, D'Alimonte I, D'Onofrio M, Nicoletti F & Caciagli F (1999) Activation of A(1) adenosine or mGlu3 metabotropic glutamate receptors enhances the release of nerve growth factor and S-100beta protein from cultured astrocytes. *Glia* **27**, 275-281
20. Edwards MM & Robinson SR (2006) TNF alpha affects the expression of GFAP and S100B: implications for Alzheimer's disease. *J Neural Transm* **113**, 1709-1715
21. Gerlach R, Demel G, Konig HG, Gross U, Prehn JH, Raabe A, Seifert V & Kogel D (2006) Active secretion of S100B from astrocytes during metabolic stress. *Neuroscience* **141**, 1697-1701
22. Nardin P, Tramontina F, Leite MC, Tramontina AC, Quincozes-Santos A, de Almeida LM, Battastini AM, Gottfried C & Goncalves CA (2007) S100B content and secretion decrease in astrocytes cultured in high-glucose medium. *Neurochem Int* **50**, 774-782
23. Goncalves D, Karl J, Leite M, Rotta L, Salbego C, Rocha E, Wofchuk S & Goncalves CA (2002) High glutamate decreases S100B secretion stimulated by serum deprivation in astrocytes. *Neuroreport* **13**, 1533-1535
24. Perrone L, Peluso G & Melone MA (2008) RAGE recycles at the plasma membrane in S100B secretory vesicles and promotes Schwann cells morphological changes. *J Cell Physiol* **217**, 60-71
25. Winningham-Major F, Staecker JL, Barger SW, Coats S & Van Eldik LJ (1989) Neurite extension and neuronal survival activities of recombinant S100 beta proteins that differ in the content and position of cysteine residues. *J Cell Biol* **109**, 3063-3071
26. Hu J, Ferreira A & Van Eldik LJ (1997) S100beta induces neuronal cell death through nitric oxide release from astrocytes. *J Neurochem* **69**, 2294-2301
27. Rothermundt M, Arolt V, Wiesmann M, Missler U, Peters M, Rudolf S & Kirchner H (2001) S-100B is increased in melancholic but not in non-melancholic major depression. *J Affect Disord* **66**, 89-93
28. Mrak RE & Griffin WS (2004) Trisomy 21 and the brain. *J Neuropathol Exp Neurol* **63**, 679-685
29. Eng LF & Ghirnikar RS (1994) GFAP and astrogliosis. *Brain Pathol* **4**, 229-237
30. Takamiya Y, Kohsaka S, Toya S, Otani M & Tsukada Y (1988) Immunohistochemical studies on the proliferation of reactive astrocytes and the expression of cytoskeletal proteins following brain injury in rats. *Brain Res* **466**, 201-210
31. Li Y, Wang J, Sheng JG, Liu L, Barger SW, Jones RA, Van Eldik LJ, Mrak RE & Griffin WS (1998) S100 beta increases levels of beta-amyloid precursor protein and its encoding mRNA in rat neuronal cultures. *J Neurochem* **71**, 1421-1428

32. Mrak RE, Sheng JG & Griffin WS (1996) Correlation of astrocytic S100 beta expression with dystrophic neurites in amyloid plaques of Alzheimer's disease. *J Neuropathol Exp Neurol* **55**, 273-279
33. Kogel D, Peters M, Konig HG, Hashemi SM, Bui NT, Arolt V, Rothermundt M & Prehn JH (2004) S100B potently activates p65/c-Rel transcriptional complexes in hippocampal neurons: Clinical implications for the role of S100B in excitotoxic brain injury. *Neuroscience* **127**, 913-920
34. Mori T, Koyama N, Arendash GW, Horikoshi-Sakuraba Y, Tan J & Town T (2010) Overexpression of human S100B exacerbates cerebral amyloidosis and gliosis in the Tg2576 mouse model of Alzheimer's disease. *Glia* **58**, 300-314
35. Pena LA, Brecher CW & Marshak DR (1995) beta-Amyloid regulates gene expression of glial trophic substance S100 beta in C6 glioma and primary astrocyte cultures. *Brain Res Mol Brain Res* **34**, 118-126
36. Hoyaux D, Decaestecker C, Heizmann CW, Vogl T, Schafer BW, Salmon I, Kiss R & Pochet R (2000) S100 proteins in Corpora amylacea from normal human brain. *Brain Res* **867**, 280-288
37. Tate-Ostroff B, Majocha RE & Marotta CA (1989) Identification of cellular and extracellular sites of amyloid precursor protein extracytoplasmic domain in normal and Alzheimer disease brains. *Proc Natl Acad Sci U S A* **86**, 745-749
38. Inoue M, Yagishita S, Itoh Y, Amano N & Matsushita M (1996) Coexistence of paired helical filaments and polyglucosan bodies in the same neuron in an autopsy case of Alzheimer's disease. *Acta Neuropathol* **92**, 511-514
39. McDonald TD, Faust PL, Bruno C, DiMauro S & Goldman JE (1993) Polyglucosan body disease simulating amyotrophic lateral sclerosis. *Neurology* **43**, 785-790
40. Tartaglia GG, Pawar AP, Campioni S, Dobson CM, Chiti F & Vendruscolo M (2008) Prediction of aggregation-prone regions in structured proteins. *J Mol Biol* **380**, 425-436
41. Koch M, Bhattacharya S, Kehl T, Gimona M, Vasak M, Chazin W, Heizmann CW, Kroneck PM & Fritz G (2007) Implications on zinc binding to S100A2. *Biochim Biophys Acta* **1773**, 457-470
42. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254
43. Morozova-Roche LA, Zurdo J, Spencer A, Noppe W, Receveur V, Archer DB, Joniau M & Dobson CM (2000) Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *J Struct Biol* **130**, 339-351
44. Malisauskas M, Zamotin V, Jass J, Noppe W, Dobson CM & Morozova-Roche LA (2003) Amyloid protofilaments from the calcium-binding protein equine lysozyme: formation of ring and linear structures depends on pH and metal ion concentration. *J Mol Biol* **330**, 879-890
45. Henry ER & Hofrichter J (1992) Singular value decomposition: application to analysis of experimental data. *Methods Enzymol* **210**, 129
46. Skerget K, Taler-Vercic A, Bavdek A, Hodnik V, Ceru S, Tusek-Znidaric M, Kumm T, Pitsi D, Pompe-Novak M, Palumaa P, et al. (2010) Interaction between oligomers of stefin B and amyloid-beta in vitro and in cells. *J Biol Chem* **285**, 3201-3210
47. Zhang J & Muthukumar M (2009) Simulations of nucleation and elongation of amyloid fibrils. *J Chem Phys* **130**, 035102
48. Kayed R, Head E, Sarsoza F, Saing T, Cotman CW, Necula M, Margol L, Wu J, Breydo L, Thompson JL, et al. (2007) Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol Neurodegener* **2**, 18
49. Barth A & Zscherp C (2002) What vibrations tell us about proteins. *Q Rev Biophys* **35**, 369-430
50. Jan A, Gokce O, Luthi-Carter R & Lashuel HA (2008) The ratio of monomeric to aggregated forms of Abeta40 and Abeta42 is an important determinant of amyloid-beta aggregation, fibrillogenesis, and toxicity. *J Biol Chem* **283**, 28176-28189
51. Hu X, Crick SL, Bu G, Frieden C, Pappu RV & Lee JM (2009) Amyloid seeds formed by cellular uptake, concentration, and aggregation of the amyloid-beta peptide. *Proc Natl Acad Sci U S A* **106**, 20324-20329
52. Gouras GK, Tsai J, Naslund J, Vincent B, Edgar M, Checler F, Greenfield JP, Haroutunian V, Buxbaum JD, Xu H, et al. (2000) Intraneuronal Abeta42 accumulation in human brain. *Am J Pathol* **156**, 15-20
53. Hofrichter J, Ross PD & Eaton WA (1974) Kinetics and mechanism of deoxyhemoglobin S gelation: a new approach to understanding sickle cell disease. *Proc Natl Acad Sci U S A* **71**, 4864-4868
54. Yan SD, Chen X, Fu J, Chen M, Zhu H, Roher A, Slattery T, Zhao L, Nagashima M, Morser J, et al. (1996) RAGE and amyloid-beta peptide neurotoxicity in Alzheimer's disease. *Nature* **382**, 685-691

55. Ostendorp T, Leclerc E, Galichet A, Koch M, Demling N, Weigle B, Heizmann CW, Kroneck PM & Fritz G (2007) Structural and functional insights into RAGE activation by multimeric S100B. *Embo J* **26**, 3868-3878
56. Leclerc E, Sturchler E & Vetter SW (2010) The S100B/RAGE Axis in Alzheimer's Disease. *Cardiovasc Psychiatry Neurol* **2010**, 539581
57. Deane R, Wu Z & Zlokovic BV (2004) RAGE (yin) versus LRP (yang) balance regulates alzheimer amyloid beta-peptide clearance through transport across the blood-brain barrier. *Stroke* **35**, 2628-2631
58. Arancio O, Zhang HP, Chen X, Lin C, Trinchese F, Puzzo D, Liu S, Hegde A, Yan SF, Stern A, et al. (2004) RAGE potentiates Abeta-induced perturbation of neuronal function in transgenic mice. *EMBO J* **23**, 4096-4105
59. Businaro R, Leone S, Fabrizi C, Sorci G, Donato R, Lauro GM & Fumagalli L (2006) S100B protects LAN-5 neuroblastoma cells against Abeta amyloid-induced neurotoxicity via RAGE engagement at low doses but increases Abeta amyloid neurotoxicity at high doses. *J Neurosci Res* **83**, 897-906
60. Chaney MO, Stine WB, Kokjohn TA, Kuo YM, Esh C, Rahman A, Luehrs DC, Schmidt AM, Stern D, Yan SD, et al. (2005) RAGE and amyloid beta interactions: atomic force microscopy and molecular modeling. *Biochim Biophys Acta* **1741**, 199-205
61. Wijesekera LC & Leigh PN (2009) Amyotrophic lateral sclerosis. *Orphanet J Rare Dis* **4**, 3
62. Hirano A, Donnenfeld H, Sasaki S & Nakano I (1984) Fine structural observations of neurofilamentous changes in amyotrophic lateral sclerosis. *J Neuropathol Exp Neurol* **43**, 461-470
63. Averback P & Langevin H (1978) Corpora amylacea of the lumbar spinal cord and peripheral nervous system. *Arch Neurol* **35**, 95-96
64. Takahashi K, Iwata K & Nakamura H (1977) Intra-axonal corpora amylacea in the CNS. *Acta Neuropathol* **37**, 165-167
65. Jellinger K & Jirasek A (1971) Neuroaxonal dystrophy in man: character and natural history. *Acta Neuropathol* **5**, Suppl 5:3-16
66. Atsumi T (1981) The ultrastructure of intramuscular nerves in amyotrophic lateral sclerosis. *Acta Neuropathol* **55**, 193-198
67. Zhang LH, Wang X, Stoltenberg M, Danscher G, Huang L & Wang ZY (2008) Abundant expression of zinc transporters in the amyloid plaques of Alzheimer's disease brain. *Brain Res Bull* **77**, 55-60
68. Bitanihirwe BK & Cunningham MG (2009) Zinc: the brain's dark horse. *Synapse* **63**, 1029-1049
69. Noy D, Solomonov I, Sinkevich O, Arad T, Kjaer K & Sagi I (2008) Zinc-amyloid beta interactions on a millisecond time-scale stabilize non-fibrillar Alzheimer-related species. *J Am Chem Soc* **130**, 1376-1383

Outlook

12. General outlook

12.1. Iron-sulfur proteins

The studies presented in this thesis resorted mainly to biophysical, biochemical and spectroscopic techniques to perform a characterization of the interplay of metal ions with protein conformation and stability. In the case of iron-sulfur (FeS) proteins, research was focused on elucidating fundamental aspects of protein folding whereas in the study of S100 proteins our experimental approach was framed in the context of S100 biology and allowed formulating hypothesis regarding pathological processes involving these proteins.

In the studies described herein, we have exploited the many spectroscopic probes available in iron-sulfur (FeS) proteins to assess protein folding properties using complementary methodologies. For the *Acidianus ambivalens* Rieske ferredoxin, the combination of fluorescence-monitored thermal unfolding, visible absorption-monitored redox titrations and visible CD-monitored pH titrations allowed us to correlate the redox status of a disulfide bridge with other properties. In the *Desulfovibrio gigas* rubredoxin and *A. ambivalens* seven iron ferredoxin, the deconvolution of temperature-resolved FT-IR spectra allowed a quantitative description of the conformational events taking place during protein unfolding. For the rubredoxin, this included a model describing protein conformation as a function of temperature. For the ferredoxin, we were able to follow the thermal unfolding of distinct structure subsets depicted as spectral components and identify distinct unfolding behaviors. The cold unfolding occurring in ferredoxin represents an uncommon event which is experimentally accessible through our deconvolution methodology.

12.2. Amyloid formation by S100 proteins

The description of amyloidogenic properties of S100 proteins and the *in vitro* amyloid cross-seeding between S100B and Aβ enclose relevant

biomedical implications and are the most relevant findings presented herein. At the moment, the studies on S100 amyloidogenesis are preliminary. Due to the fully uncharacterized amyloidogenic properties of the S100 proteins in this study, we have restricted our studies to acidic pH. Although these conditions do not allow prompt assumption that S100 amyloid formation occurs *in vivo*, it nevertheless establishes the proof of concept for future studies.

S100A6 [1], S100A12 [2] and S100B [3] are overexpressed in Alzheimer's disease accumulating in the synaptic space as response to Aβ deposition. In this milieu there are high concentrations of metal ions which may facilitate amyloid formation by S100 proteins (*e. g.* Cu^{2+}-induced S100B amyloidogenesis). Based on this possibility, we assessed the cytotoxicity of S100 amyloid species. Indeed, we detected cytotoxicity arising from S100A6 oligomers as well as S100A12 and S100B mature amyloids, an indication of the pathological relevance of these species if they form *in vivo*. The molecular basis for cytotoxicity was not investigated, but membrane disruption is a common mechanism among different amyloid species. Specifically for S100 proteins, the transduction of the extracellular protein accumulation into intracellular events may involve the Receptor for Advanced Glycation Endproducts (RAGE), a known S100 ligand [4]. Interestingly, RAGE also binds Aβ oligomers and fibrils, enclosing the possibility for RAGE-Aβ-S100 cross-talk in cytotoxic events.

Based on the abundant secretion of S100A12 and S100B and, to a lesser extent, S100A6, in the synaptic metal-rich milieu where Aβ oligomeric and fibrillar structure accumulate, we have hypothesized that S100-Aβ cross-seeding could be instrumental in disease progression. If S100 proteins are amyloidogenic at physiological conditions, this interaction could constitute a mechanism facilitating onset of fibrillization. The imbalancement in metal ion homeostasis occurring in Alzheimer's disease could act as a further factor

facilitating amyloidogenesis. Further, metal ions are known to influence Aβ amyloidogenesis [5]. Our results on the modulation of S100 amyloidogenic pathways by metal ions support this hypothesis.

Subsequent studies on the biomedical implications of S100 amyloid formation should establish the conditions where S100 proteins are amyloidogenic and determine if amyloid formation is accessible under physiological conditions. This knowledge will allow formulating hypothesis on the conditions where the amyloidogenic process of S100 proteins and Aβ can be correlated. An identical approach may apply to the involvement of S100A6 in ALS [6]. Overexpression of this mainly intracellular protein in a disease characterized by aggregation of the also intracellular protein superoxide dismutase 1 (SOD1) may be analyzed in an analogous manner. The determination of the potential implication of S100 proteins in neurodegenerative diseases will build on the knowledge of the molecular basis of disease and be useful in the design of novel prophylactic therapeutic strategies.

12.3. References

1. Boom A, Pochet R, Authelet M, Pradier L, Borghgraef P, Van Leuven F, Heizmann CW & Brion JP (2004) Astrocytic calcium/zinc binding protein S100A6 over expression in Alzheimer's disease and in PS1/APP transgenic mice models. Biochim Biophys Acta 1742, 161-168
2. Shepherd CE, Goyette J, Utter V, Rahimi F, Yang Z, Geczy CL & Halliday GM (2006) Inflammatory S100A9 and S100A12 proteins in Alzheimer's disease. Neurobiol Aging 27, 1554-1563
3. Mrak RE, Sheng JG & Griffin WS (1996) Correlation of astrocytic S100 beta expression with dystrophic neurites in amyloid plaques of Alzheimer's disease. J Neuropathol Exp Neurol 55, 273-279
4. Leclerc E, Sturchler E & Vetter SW (2010) The S100B/RAGE Axis in Alzheimer's Disease. Cardiovasc Psychiatry Neurol 2010, 539581
5. Meloni G, Sonois V, Delaine T, Guilloreau L, Gillet A, Teissie J, Faller P & Vasak M (2008) Metal swap between Zn7-metallothionein-3 and amyloid-beta-Cu protects against amyloid-beta toxicity. Nat Chem Biol 4, 366-372
6. Hoyaux D, Boom A, Van den Bosch L, Belot N, Martin JJ, Heizmann CW, Kiss R & Pochet R (2002) S100A6 overexpression within astrocytes associated with impaired axons from both ALS mouse model and human patients. J Neuropathol Exp Neurol 61, 736-744